

us-10-054-579-3.rnppm

Mon Dec 1 12:41:17 2003

```

1261 TCACAGCCCACTCAGCAGCCCGGGTACCCCTCACCCTCAGCAAGGGGAGTCCC 1320
1141 CTCCTCCACCCCAAGGAGCAGCTTCCACAGCCCAAGAGAGACCCCGGTGCGCAGCCC 1200
1321 CTCCTCCACCCCAAGGAGCAGCTTCCACAGCCCAAGAGAGAGCCCGGTGCGCAGCCC 1380
1201 AACCCAGCCCGCTCAGCAGCCCGGTGAGGGGTGCTGAGGGGCGCGCTCAAC 1260
1381 AACCCAGCCCGCTCAGCAGCCCGGTGAGGGGTGCTGAGGGGCGCGCTCAAC 1440
1261 TCCATCAAGAGCAGCTTTCCTGAGTCCACCGCTTCCACCGCGGAATCTCAAGTTCCG 1320
1441 TCCATCAAGAGCAGCTTTCCTGAGTCCACCGCTTCCACCGCGGAATCTCAAGTTCCG 1500
1321 AGCCGAGGAGATGTCACCACTGACACAGAGTGTCTCCAGAGCTGCGGAGAGTCC 1380
1501 AGCCGAGGAGATGTCACCACTGACACAGAGTGTCTCCAGAGCTGCGGAGAGTCC 1560
1381 TGGTTTGGGAACTTCATCAGCTTGGAGAGGAGGAGAGATCTTGTGTATCAAGAGC 1440
1561 TGGTTTGGGAACTTCATCAGCTTGGAGAGGAGGAGAGATCTTGTGTATCAAGAGC 1620
1441 AAACCTCTGAGTCCATCAAGGCTGACATGTCGACGCTTCTGTCGATTCCAGTCTC 1500
1621 AAACCTCTGAGTCCATCAAGGCTGACATGTCGACGCTTCTGTCGATTCCAGTCTC 1680
1501 AGCCAGAGCTATCTCCCAAGAGGATTCCTCCGCGGAGTACAGGCCACGCGGGGCCA 1560
1681 AGCCAGAGCTATCTCCCAAGAGGATTCCTCCGCGGAGTACAGGCCACGCGGGGCCA 1740
1561 GCGGTGTTCCAGAGCGGCTCAAGTTCAGGTTGATATCACTACACGAGGAGTGGGAG 1620
1741 GCGGTGTTCCAGAGCGGCTCAAGTTCAGGTTGATATCACTACACGAGGAGTGGGAG 1800
1621 GCGAGAGAGGAGAGCGGCTTACTCCGTCACCTTCACTCCCTGCTCAGGCCCGCCGCT 1680
1801 GCGAGAGAGGAGAGCGGCTTACTCCGTCACCTTCACTCCCTGCTCAGGCCCGCCGCT 1860
1681 GCGTTCAAGAGGTTGGTGGAGACCATCAGGCCAGCTGCTGAGCAGCAGCCCGCT 1740
1861 GCGTTCAAGAGGTTGGTGGAGACCATCAGGCCAGCTGCTGAGCAGCAGCCCGCT 1920
1741 GCGGCCAGCAGCTTGTGAGCAGCAGCTTGTATGGAATGATGAGCGGCGCTTCC 1800
1921 GCGGCCAGCAGCTTGTGAGCAGCAGCTTGTATGGAATGATGAGCGGCGCTTCC 1980
1801 AAATGTGGAATTTCCGAAAGTTAA 1827
1981 AAATGTGGAATTTATCCGAAAGTTAA 2007

```

RESULT 3

US-10-288-798-42

```

; Sequence 42, Application US/10288798
; GENERAL INFORMATION:
; APPLICANT: BANDMAN, Olga; NGUYEN, Daniel B.;
; APPLICANT: WALIA, Narinder K.; HAFALIA, April J.A.;
; APPLICANT: YAO, Monique G.; GANDHI, Ameena R.;
; APPLICANT: GURURAJAN, Rajagopal; DING, Li;
; APPLICANT: PATTERSON, Chandra; YUE, Henry;
; APPLICANT: BAUGHN, Mariah R.; TRIBOULEY, Catherine M.;
; APPLICANT: THORNTON, Michael; ELLIOTT, Vicki S.;
; APPLICANT: LU, Yan; ISON, Craig H.;
; APPLICANT: AU-YOUNG, Janice; TANG, Y. Tom;
; APPLICANT: AZIMZAI, Valda; BURRILL, John D.;
; APPLICANT: MARCUS, Gregory A.; ZINGLER, Kurt A.;
; APPLICANT: LU, Dying Aina M.; LAL, Preethi G.;
; APPLICANT: RAMKUMAR, Jayalaxmi; WARREN, Bridget A.;
; APPLICANT: KEARNEY, Liam; POLICKY, Jennifer L.;
; APPLICANT: THANGAVELU, Kavitha; BURFORD, Neil
; TITLE OF INVENTION: HUMAN KINASES
; REFERENCE: FI-0209 USA

```

```

; CURRENT APPLICATION NUMBER: US/10/288,798
; CURRENT FILING DATE: 2002-11-01
; PRIOR APPLICATION NUMBER: PCT/US01/27219
; PRIOR FILING DATE: 2001-08-31
; PRIOR APPLICATION NUMBER: US 60/240,542
; PRIOR FILING DATE: 2000-10-13
; PRIOR APPLICATION NUMBER: US 60/238,389
; PRIOR FILING DATE: 2000-10-06
; PRIOR APPLICATION NUMBER: US 60/236,499
; PRIOR FILING DATE: 2000-09-29
; PRIOR APPLICATION NUMBER: US 60/234,902
; PRIOR FILING DATE: 2000-09-22
; PRIOR APPLICATION NUMBER: US 60/232,654
; PRIOR FILING DATE: 2000-09-14
; PRIOR APPLICATION NUMBER: US 60/231,357
; PRIOR FILING DATE: 2000-09-08
; PRIOR APPLICATION NUMBER: US 60/229,873
; PRIOR FILING DATE: 2000-08-31
; NUMBER OF SEQ ID NOS: 48
; SOFTWARE: PERL Program
; SEQ ID NO 42
; LENGTH: 2647
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: Incyte ID No: 4022651CB1
; US-10-288-798-42

Query Match 100.0%; Score 1827; DB 49; Length 2647;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1827; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGAAGTGGAGCGGGAGATCGGATCTCTGAAGCTCATTGAGCACCCTCCAGCTCTTAAG 60
DB 169 ATGAAGTGGAGCGGGAGATCGGATCTCTGAAGCTCATTGAGCACCCTCCAGCTCTTAAG 228
QY 61 CTGACGAGCTTTATGAAACAAATAATTTGTACCTGGTGTAGAACACCTGTGAGT 120
DB 229 CTGACGAGCTTTATGAAACAAATAATTTGTACCTGGTGTAGAACACCTGTGAGT 288
QY 121 GGTGAGCTCTTGGACTACCTGTTGAGAGGGAGGCTGACGCTTAAGGAGGCTCGGAG 180
DB 289 GGTGAGCTCTTGGACTACCTGTTGAGAGGGAGGCTGACGCTTAAGGAGGCTCGGAG 348
QY 181 TTCTTCGGGAGATCATCTCTGCGTGGACTCTGTCACAGCCACTCCATATGCCAGG 240
DB 349 TTCTTCGGGAGATCATCTCTGCGTGGACTCTGTCACAGCCACTCCATATGCCAGG 408
QY 241 GATCTGAAACCTGAAACCTCTCTGCTGGAGAGAGAACAAATCCGCATCGCAGCTTT 300
DB 409 GATCTGAAACCTGAAACCTCTCTGCTGGAGAGAGAACAAATCCGCATCGCAGCTTT 468
QY 301 GGCAATGGCGTCCCTGCGAGGTGGCGAGAGCTGTTGAGAGACCACTGTTGGTCCCCC 360
DB 469 GGCAATGGCGTCCCTGCGAGGTGGCGAGAGCTGTTGAGAGACCACTGTTGGTCCCCC 528
QY 361 TACGCTGCCCGAGGAGTATCCGCGGGGAGAGATGACGCGCGGAGAGGCGGAGCGTGG 420
DB 529 TACGCTGCCCGAGGAGTATCCGCGGGGAGAGATGACGCGCGGAGAGGCGGAGCGTGG 588
QY 421 AGCTGCGGCGTCACTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 480
DB 589 AGCTGCGGCGTCACTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 648
QY 481 TTGCGACAGCTGCTGAGAGAGTGAAGCGGGCGTGTTCACATCCGCACTTTATCCCG 540
DB 649 TTGCGACAGCTGCTGAGAGAGTGAAGCGGGCGTGTTCACATCCGCACTTTATCCCG 708
QY 541 CCGGACTGCGAGCTGCTGCTGCGGCGATGATCGAGGTGGAGCGCGCAGCGCCCTCAGG 600
DB 709 CCGGACTGCGAGCTGCTGCTGCGGCGATGATCGAGGTGGAGCGCGCAGCGCCCTCAGG 768

```

Qy	601	CTRAGACATTTTCAGAAAACATATGGTATATAGGGGGCAGAAATGAGCCCGAAACAGAG	660
Db	769	CTAGAGCATTTCAGAAAACATATGGTATATAGGGGGCAAGAATGAGCCCGAAACAGAG	828
Qy	661	CAGCCATTCTCGCAAGGTGCAGATCCGCTCGCTGCCAGCCTGGAGGACATTCGACCCC	720
Db	829	CAGCCATTCTCGCAAGGTGCAGATCCGCTCGCTGCCAGCCTGGAGGACATTCGACCCC	888
Qy	721	GACGTGCTGAGCAGCATGCACTCACTGGGTGTCTTCGAGACCGCAACAAAGTGTCTGCAG	780
Db	889	GACGTGCTGAGCAGCATGCACTCACTGGGTGTCTTCGAGACCGCAACAAAGTGTCTGCAG	948
Qy	781	GACCTGCTGCCAGAGGAGAAACAGAGAAAGATTTACTTCTCTCTCTGGACCCG	840
Db	949	GACCTGCTGCCAGAGGAGAAACAGAGAAAGATTTACTTCTCTCTCTGGACCCG	1008
Qy	841	AAAGAAAGGTACCCGAGCCAGGAGATGAGGACCTGCCCCCGCGAAACGAGATAGACCT	900
Db	1009	AAAGAAAGGTACCCGAGCCAGGAGATGAGGACCTGCCCCCGCGAAACGAGATAGACCT	1068
Qy	901	CCCCGAAGCGTGTGCACTCCCGATGCTGAAACGGCAACGGCAAGCGCGGCCAGAACGC	960
Db	1069	CCCCGAAGCGTGTGCACTCCCGATGCTGAAACGGCAACGGCAAGCGCGGCCAGAACGC	1128
Qy	961	AAATTCATGGAGTGTCTCAGCGTGACCGACGGCGGCTCCCGGTGCTCGCGCGCGGGCC	1020
Db	1129	AAATTCATGGAGTGTCTCAGCGTGACCGACGGCGGCTCCCGGTGCTCGCGCGCGGGCC	1188
Qy	1021	ATTGAGATGGCCCAACAGCAGGCAAGAGTCTCGTTCATCAGCGGTGCTCTCTCAGAGCTT	1080
Db	1189	ATTGAGATGGCCCAACAGCAGGCAAGAGTCTCGTTCATCAGCGGTGCTCTCTCAGAGCTT	1248
Qy	1081	TCCACACGCCCACTCAGCAGCCCCCGGTGACCCCTCACCCCTCACCAAGGGGCACTCCC	1140
Db	1249	TCCACACGCCCACTCAGCAGCCCCCGGTGACCCCTCACCCCTCACCAAGGGGCACTCCC	1308
Qy	1141	CTCCCAACCCCAAGGGGACACCTGTGTCCAACGCAAGAGAGACCCGGCTGGGCAAGCC	1200
Db	1309	CTCCCAACCCCAAGGGGACACCTGTGTCCAACGCAAGAGAGACCCGGCTGGGCAAGCC	1368
Qy	1201	AACCCACGCCCGCTCAGCCCCCAGCGTCGAGGGGTGCCCTGAGGGCGCGGCTCAAC	1260
Db	1369	AACCCACGCCCGCTCAGCCCCCAGCGTCGAGGGGTGCCCTGAGGGCGCGGCTCAAC	1428
Qy	1361	TCCATCAAGAACAGCTTTCTGGGCTCACCCCGTCTTCAACCGCCGGAATCTCAAGTTCCG	1320
Db	1429	TCCATCAAGAACAGCTTTCTGGGCTCACCCCGTCTTCAACCGCCGGAATCTCAAGTTCCG	1488
Qy	1321	ACGCCGGAGGAGTGTCCAACTTCACACCCAGAGTCGTCCCGAGAGCTGGCGAGAGTCC	1380
Db	1489	ACGCCGGAGGAGTGTCCAACTTCACACCCAGAGTCGTCCCGAGAGCTGGCGAGAGTCC	1548
Qy	1381	TGGTTTGGGAACTTCATCAGCCTCGAGAGGAGGAGCAGATCTCTGTGGTTCATCAAGAC	1440
Db	1549	TGGTTTGGGAACTTCATCAGCCTCGAGAGGAGGAGCAGATCTCTGTGGTTCATCAAGAC	1508
Qy	1441	AAACCTCTGAGCTCCATCAAGGCTGACATCGTGCACGCTTCTGTGATTTCCAGTCTC	1500
Db	1609	AAACCTCTGAGCTCCATCAAGGCTGACATCGTGCACGCTTCTGTGATTTCCAGTCTC	1568
Qy	1501	AGCCACAGCTCATCTCCAAACAGACTTCCGGGCGAGTACAAGGCCACGGGGGGGCCA	1560
Db	1669	AGCCACAGCTCATCTCCAAACAGACTTCCGGGCGAGTACAAGGCCACGGGGGGGCCA	1728
Qy	1561	GCCGTGTTCCAGAACCCGGTCAAGTTCAGGTTGATATCACTTACAAGGAGGTGGGAG	1620
Db	1729	GCCGTGTTCCAGAACCCGGTCAAGTTCAGGTTGATATCACTTACAAGGAGGTGGGAG	1788
Qy	1621	GGCAGAGAGGAGACGGCATCTACTCGTCACTTCAACCTGCTCTCAGGCCCCAGCCGT	1680
Db	1789	GGCAGAGAGGAGACGGCATCTACTCGTCACTTCAACCTGCTCTCAGGCCCCAGCCGT	1848
Qy	1681	CGCTTCAAGAGGGGTGGAGACCATCCAGGCCACAGCTGTGTGAGCACACAGACCCGCT	1740

```

Db      1849  CGCTTCAAGAGGGTGTGAGACCATCCAGGCCCGCTCTGAGACACACGACCCGCTT 1908
QY      1741  GGGCCCCAGCACCTTGTGCAGACACCACTAATCTGTATGGAAATGATGACGGGGCGGCTTTTC 1800
Db      1909  GCGGCCAGCACCTTGTGCAGACACCACTAATCTGTATGGAAATGATGACGGGGCGGCTTTTC 1968
QY      1801  AATGTGGAATTATCCGAAAGTTAA 1827
Db      1969  AATGTGGAATTATCCGAAAGTTAA 1995

RESULT 4
US-10-362-892-42
; Sequence 42, Application US/10362892
; GENERAL INFORMATION:
; APPLICANT: INCYTE GENOMICS, INC.; BANDMAN, Olga
; APPLICANT: NGUYEN, Damiel B.; WALIA, Narinder K.
; APPLICANT: HAFALIA, April J.A.; YAO, Monique G.
; APPLICANT: GANDHI, Ameena R.; GURURAJAN, Rafagopal
; APPLICANT: DING, Li; PATTERSON, Chandra S.
; APPLICANT: YUE, Henry; BAUGHN, Mariah R.
; APPLICANT: TRIBOULEY, Catherine M.; THORNTON, Michael B.
; APPLICANT: ELLIOTT, Vicki S.; LU, Yan
; APPLICANT: TANG, Y. Tom; AZIMZAI, Yalda
; APPLICANT: BURRILL, John D.; MARCUS, Gregory A.
; APPLICANT: ZINGLER, Kurt A.; LU, Dying Aina M.
; APPLICANT: LAL, Preeti G.; RAMKUMAR, Jayalaxmi
; APPLICANT: WARREN, Bridget A.; KEARNEY, Liam
; APPLICANT: POLICKY, Jennifer L.; THANGAVELOU, Kavitha
; APPLICANT: BURFORD, Neil
; TITLE OF INVENTION: HUMAN KINASES
; FILE REFERENCE: PF-0209 USN
; CURRENT APPLICATION NUMBER: US/10/362,892
; CURRENT FILING DATE: 2003-02-25
; PRIOR APPLICATION NUMBER: PCT/US01/27219
; PRIOR FILING DATE: 2001-08-31
; PRIOR APPLICATION NUMBER: US 60/229,873
; PRIOR FILING DATE: 2000-08-31
; PRIOR APPLICATION NUMBER: US 60/231,357
; PRIOR FILING DATE: 2000-09-08
; PRIOR APPLICATION NUMBER: US 60/232,654
; PRIOR FILING DATE: 2000-09-14
; PRIOR APPLICATION NUMBER: US 60/234,902
; PRIOR FILING DATE: 2000-09-22
; PRIOR APPLICATION NUMBER: US 60/236,499
; PRIOR FILING DATE: 2000-09-29
; PRIOR APPLICATION NUMBER: US 60/238,389
; PRIOR FILING DATE: 2000-10-06
; PRIOR APPLICATION NUMBER: US 60/240,542
; PRIOR FILING DATE: 2000-10-13
; NUMBER OF SEQ ID NOS: 48
; SOFTWARE: PERL Program
; SEQ ID NO 42
; LENGTH: 2647
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: Indbyte ID No: 4022651CB1
US-10-362-892-42

Query Match      100.0%; Score 1827; DB 51; Length 2647;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1827; Conservative 0; Mismatches 0; Gaps 0

QY      1      ATGAAGTGTGAGCGGGAGATCGCGATCTCTGAAGCTCATTTGAGCACCCCGGCTCTTAAG 60
Db      169  ATGAGGTGTGAGCGGGAGATCGCGATCTCTGAGCTCATTTGAGCACCCCGGCTCTTAAG 228
QY      61  CTGACGACGTTTATGAAAAACAAAAATATTTGTACTGGTGTAGAAACACGTGTCAAGT 120

```


Query No.	Score	Query			DB ID	Description
		Match	Length	%		
1	763.2	38.0	1201	91	BQ538014	AL
2	639.4	31.9	769	13	BQ442940	U
3	633.2	31.5	791	13	BQ611869	U
4	601.4	30.0	732	13	BQ443783	U

RESULT 1
 AL538014
 LOCUS
 DEFINITION
 ACCESSION
 VERSION
 KEYWORDS
 SOURCE
 ORGANISM
 Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 1 (bases 1 to 1201)
 Li, W.B., Gruber, C., Jessee, J. and Polayes, D.
 Full-length cDNA libraries and normalization
 Unpublished
 On Feb 13, 2001 this sequence version replaced gi:12801507.
 Contact: Genoscope
 Genoscope - Centre National de Sequençage
 BP 191 91006 EVRY cedex - France
 Email: seqes@genoscope.cns.fr, Web : www.genoscope.cns.fr
 Library was constructed by Life Technologies, a division of
 Invitrogen. This sequence belongs to sequence cluster 6094.r For
 more information about this cluster, see
 http://www.genoscope.cns.fr/
 cgi-bin/cluster.cgi?seq=CS0DF029AB11QPl&cluster=6094.r. Contact :
 Feng Liang Email : fliang@lifetech.com URL :

<http://fulllength.invitrogen.com/> InvitroGen Corporation 1600
Faraday Avenue Genoscope sequence ID : CS0DF029AB11QPL.

FEATURES
source

```

source
location/Qualifiers
1..1201
/organism="Homo sapiens"
/mol_type="mrna"
/db_xref="taxon:9606"
/clone="CSODF029YC21"
/tissue_type="FETAL BRAIN"
/dev_stage="fetal"
/clone_lib="Homo sapiens FETAL BRAIN"
/note="Organ: brain; Vector: pcwvSPORT 6; 1st strand cDNA
was primed with a NotI-oligo(dT) primer. Five prime end
enriched, double-strand cDNA was digested with Not I and
cloned into the Not I and EcoRV sites of the pcwvSPORT 6
vector. Library was not normalized."

BASE COUNT      273 a 282 c 364 g 204 t 78 others
ORIGIN

Query Match
Best Local Similarity 38.0%; Score 763.2; DB 9; Length 1201;
Matches 824; Conservative 10; Mismatches 18; Indels 6; Gaps 5;

QY 1 ATGACATCGACGGGGAAGGACGGCGCGCGCAGCAGCGCAGTATGTTGGCCCTACCGG 60
DB |||
87 ATGACATCGACGGGGAAGGACGGCGCGCGCA-CACGCGCAGTATGTTGGCCCTACCGG 145
QY 61 CTGGAGAAGACGCTGGGCAAGGGGCGACAGGTCTGTGTGAAGCTGGGGTTCACTGCGTC 120
DB |||
146 CTGGAGAAGACGCTGGGCAAGGGGCGACAGGTCTGTGTGAAGCTGGGGTTCACTGCGTC 205
QY 121 ACCTGCCAAGAGTGGCCATCAAGATCGTCAACCGTGAGAAGCTCACGAGTCGGTCTG 180
DB |||
206 ACCTGCCAAGAGTGGCCATCAAGATCGTCAACCGTGAGAAGCTCACGAGTCGGTCTG 265
QY 181 ATGAGGTGGAGCGGGAGATCGGATCTCTGAAGTCTATTGAGCACCCCGCAGTCTTAAAG 240
DB |||
266 ATGAAGGTGGAGCGGGAGATCGGATCTCTGAAGTCTATTGAGCACCCCGCAGTCTTAAAG 325
QY 241 CTGCACGACGTTTATGAACACAAAAATATTTGTACCTGGTGTCTAGAACACGTGTCAAGT 300
DB |||
326 CTGCACGACGTTTATGAACACAAAAATATTTGTACCTGGTGTCTAGAACACGTGTCAAGT 385
QY 301 GGTGAGCTCTTCGACTACCTGGTGAAGAGGGGAGGTGACGCTTAAGGAGGCTCGGAG 360
DB |||
386 GGTGAGCTCTTCGACTACCTGGTGAAGAGGGGAGGTGACGCTTAAGGAGGCTCGGAG 445
QY 361 TTCTTCGGGAGATCATCTCTGCGCTGGAATTCTGCGCACGCCATCTCCATATGCCACAGG 420
DB |||
446 TTCTTCGGGAGATCATCTCTGCGCTGGAATTCTGCGCACGCCATCTCCATATGCCACAGG 505
QY 421 GATCTGAAACCTGAAAACTCTCTGCGACGAGAAGAACAACTCCGCATCGCAGACTTTT 480
DB |||
506 GATCTGAAACCTGAAAACTCTCTGCGACGAGAAGAACAACTCCGCATCGCAGACTTTT 565
QY 481 GGCATGGGCTCCCTGCGAGGTGGCGACAGGCTGTTGGAGACGAGCTGTGGTCCCCCCCAC 540
DB |||
566 GGCATGGGCTCCCTGCGAGGTGGCGACAGGCTGTTGGAGACGAGCTGTGGTCCCCCCCAC 625
QY 541 TAGCCCTCCCCGAGGTGATCCGGGGGGAAGATGACGCGCGGAAGCGGACGTGTGG 600
DB |||
626 TAGCCCTCCCCGAGGTGATCCGGGGGGAAGATGACGCGCGGAAGCGGACGTGTGG 684
QY 601 AGCTGCGGCGTCACTCTGTTGCGCTTGTCTGCTGGGGGCTCTGCCCTTCACCATGACAC 660
DB |||
685 AGCTGCGGCGTCACTCTGTTGCGCTTGTCTGCTGGGGGCTCTGCCCTTCACCATGACAC 744
QY 661 TTTCGACAGTCTGAGAAAGGTGAACCGGGGCGTGTTCACATGCCGACATTTATCCCG 720
DB |||
745 TTTCGACA-CTGCTGGAGAGGTGAACCGGGGCGTGTTCACATGCCGACATTTATCCCG 803
QY 721 CCCGACTGCCAGAGTCTGTACGGGCGATGATCAGGTGGAACCGCGACCGCGCTCAG 780
DB |||

```

Db	804	CCGACTGCAGAACTCTGCTACGGGGCATGATCGAGGTGGACGCCGCACGCCGCCTCAGC	863
Qy	781	CTAGAGCACATTTCAGAAACACATATGTTATATAGGGGCAAGAATGAGCCGGAACACAG	840
Db	864	CTAGAGCACATTTCAGAAACACATATGKWTATAGGGGSA--AAATGAGCCSAAACAG	921
Qy	841	CAGCCCATTCCTCGCAAG	858
Db	922	CAG--CCATTCTCTGCAAG	938

RESULT 2
BO442940

LOCUS	BQ442940	769 bp	linear	EST 29-MAY-2002
DEFINITION	UI-M-EVO-bxf-o-02-0-UI.x1 NIH BMAP_EVO Mus musculus cDNA clone IMAGE:5707609 5', mRNA sequence.			
ACCESSION	BQ442940			
VERSION	BQ442940.1	GI:21246052		
KEYWORDS	EST.			
SOURCE	Mus musculus (house mouse)			
ORGANISM	Mus musculus			
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.			
REFERENCE	1 (bases 1 to 769)			
AUTHORS	NIH-MGC http://mgc.nci.nih.gov/ .			
TITLE	National Institutes of Health, Mammalian Gene Collection (MGC)			
JOURNAL	Unpublished			
COMMENT	Contact: Robert Strausberg, Ph.D.			

Email: gapus-remail@linl.gov
 Tissue Procurement: Dr. James Lin, University of Iowa
 cDNA Library preparation: Dr. M. Bento Soares, University of Iowa
 cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
 DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
 Clone Distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>

This clone was contributed by the Brain Molecular Anatomy Project (BMAP)

```

FEATURES
    Seq primer: pyx-5,
    Location/Qualifiers
        1..769
            /organism="Mus musculus"
            /mol_type="mRNA"
            /strain="C57BL/6"
            /db_xref="taxon:10090"
            /clone="IMAGE:5707609"
            /tissue_type="whole brain"
            /dev_stage="embryo 15.5 dpc"
            /lab_host="DH10B (T1 phage resistant)"

```

195 a 241 c 202 q 141 t

BRASE COUNT

Query Match 31.9%; Score 639.4; DB 13; Length 769;
Best Local Similarity 89.5%; Pred. No. 1.9e-139;
Matches 688; Conservative 0; Mismatches 81; Indels 0; Gaps 0;
901 GACGTGTTGGACAGCATGCACCTCACTGGGTGCTTCCGAGACCGCAACAAGTGTCTGCAG 960

Db 481 GCCTTGCTGGGGCTCTGCTTTTGTATGATGACAACTCGGCGAGTGTCTGGAGAG 540
 Qy 582 GTGAAGCGGGCGTGTTCACATCCGACATTTATCCGCGCCGAGTCCGACAGTGTCTGA 741
 Db 541 GTCAAGCGTGTGTGTTCACATCCGACATTTATCCGCGCCGAGTCCGACAGTGTCTCTG 600
 Qy 742 CGGGCGCATGATCGAGTGTGGAGCGCGCACCGCGCTCACGCTAGAGCACATTCAGAAACAC 801
 Db 601 CGTGCGCATGATGAGTGTGGAGCGCGCTCACGCTAGAGCACATTCAGAAACAC 660
 Qy 802 ATATGTTATAGGGGCAAGAAATGAGCGCGCAACAGAGAGCGCCCATTCCTCGCAAGGTG 861
 Db 661 ATATGTTATAGTGTGGCAAGAAATGAGCGCGCAACAGAGAGCGCCCATTCCTCGCAAGGTG 720
 Qy 862 CAGATCGCTCGTCCGAGCTGGAGACATCACCCGAGCTGCTGGAGAGCATGCAC 921
 Db 721 CAGATCGCTCACATCCGAGCTGGAGACATTCACCCGAGCTGCTGGAGAGCATGCAC 780
 Qy 922 TCACTGGGCTG 932
 Db 781 TCACTGGGCTG 791

RESULT 4
 BQ443783
 LOCUS
 DEFINITION
 BQ443783 732 bp mRNA linear EST 29-MAY-2002
 IMAGE:5708522 5', mRNA sequence.
 BQ443783
 EST.
 SOURCE
 Mus musculus (house mouse)
 ORGANISM
 Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 1 (bases 1 to 732)
 NIH-MGC http://mgi.nci.nih.gov/
 National Institutes of Health, Mammalian Gene Collection (MGC)
 Unpublished
 Contact: Robert Strausberg, Ph.D.
 Email: csapbs-remail.nih.gov
 Tissue Procurement: Dr. James Lin, University of Iowa
 cDNA Library Preparation: Dr. M. Bento Soares, University of Iowa
 cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
 DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
 Clone Distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LNL at:
 http://image.llnl.gov
 This clone was contributed by the Brain Molecular Anatomy Project
 (BMAP)

Seq primer: pyx-5.
 Location/Qualifiers
 1. 732
 /organism="Mus musculus"
 /mol_type="mRNA"
 /strain="C57BL/6"
 /db_xref="taxon:10090"
 /clone="IMAGE:5708522"
 /tissue_type="whole brain"
 /dev_stage="embryo 15.5 dpc"
 /lab_host="DH10B (T1 phage resistant)"
 /clone_lib="NIH-BMAP_EWO"
 /note="Organ: brain; Vector: pyx-Asc; Site 1: EcoR I;
 Site 2: Not I; The library was constructed according to
 Bonaldi, Lennon and Soares, Genome Research, 6:791-806,
 1996. Denatured mRNA was size fractionated on a 1% agarose
 gel. First strand cDNA synthesis was primed with an
 oligo-dT primer containing a Not I site. Double stranded
 cDNA was size selected according to mRNA size fraction,
 ligated with EcoR I adaptor, digested with Not I, and then
 cloned directionally into pyx-Asc vector. The library tag
 sequence located between the Not I site and the polyA tail
 , is GTGCGTGGAA. This library was created for the

University of Iowa Mouse Brain Molecular Anatomy Project
 (BMAP): 'Gene Discovery in the Developing Mouse Nervous
 System', supported by National Institutes of Mental Health
 (NIMH), Hemin Chin, Ph.D., program coordinator."

BASE COUNT 178 a 240 c 188 g 125 t 1 others

Query Match 30.0%; Score 601.4; DB 13; Length 732;
 Best Local Similarity 88.8%; Pred. No. 3.9e-121;
 Matches 650; Conservative 0; Mismatches 82; Indels 0; Gaps 0;

Qy 817 GGCAGAAATGAGCGCGCAACAGAGAGCGCCATTCCTCGCAAGGTGAGATCCGCTCGCTG 876
 Db 1 GGCAGAAATGAGCGCGCGCAACAGAGAGCGCCATTCCTCGCAAGGTGAGATCCGCTCACTA 60
 Qy 877 CCAGAGCTGGAGGACATCGACCCCGACGTGCTGGAGAGAGAGAGAGAGAGAGAGAG 936
 Db 61 CCAGAGCTGGAGGACATTCGACCTGATGTGTGGAGAGAGAGAGAGAGAGAGAGAG 120
 Qy 937 CGAGACCGCAACAGCTGTCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 996
 Db 121 CGAGACCGCAACAGCTGTCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 180
 Qy 997 ATTACTCTCTCTCTCTGACCGGAAAGAGTATCCCGAGCGAGAGAGAGAGAGAGAGAG 1056
 Db 181 ATTATTTCTCTCTCTGATCGGAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 240
 Qy 1057 CCCCCCGGAGAGAGATAGACCTCTCCCGGAGCGTGTGGAGTCTCCCGGAGTGTGAACCGG 1116
 Db 241 CCCCCCGGAGAGATAGATAGACCTCTCCCGGAGCGTGTGGAGTCTCCCGGAGTGTGAACCGG 300
 Qy 1117 CAGCGCAAGCGGGCGCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1176
 Db 301 CATGGCAAGCGCGGAG 360
 Qy 1177 TCCCGGCTCTCTCTCTGCGCGCGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1236
 Db 361 TCCCGGAGTCTCTCTCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 420
 Qy 1237 ATCAGCGGTCTCTCTCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1296
 Db 421 ATCAGTGTGTCTCTCTCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 480
 Qy 1297 CACCCCTTACCAAGGGGAGTGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1356
 Db 481 CACCCCTTACCAAGGGGAGTGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 540
 Qy 1357 AAGGAGAGCGCGGTGCGCACGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1416
 Db 541 AAGGAGAGCGCGGTGCGCACGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 600
 Qy 1417 GTGCGCTGAGGGCGCGGTCTCAACTCAATCAAGAGAGAGAGAGAGAGAGAGAGAGAG 1476
 Db 601 GTGCGCTGAGGGCGCGGTCTCAACTCAATCAAGAGAGAGAGAGAGAGAGAGAGAGAG 660
 Qy 1477 CACCGCGGAAATGCAAGTTCGAGCGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1536
 Db 661 CACCGCGGAAATGCAAGTTCGAGCGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 720
 Qy 1537 TCCCCAGAGCTG 1548
 Db 721 TCTCCAGAGCTG 732

RESULT 5
 CA315127
 LOCUS
 DEFINITION
 CA315127 723 bp mRNA linear EST 26-NOV-2002
 IMAGE: 6810176 5', mRNA sequence.
 ACCESSION
 CA315127
 VERSION
 CA315127.1 GI:24533251
 KEYWORDS
 EST.
 Mus musculus (house mouse)

(BMAP)

Seq primer: pyx-5.

Location/Qualifiers

ES

source

1. 712

/organism="Mus musculus"

/mol_type="mRNA"

/strain="C57BL/6"

/db_xref="taxon:10090"

/clone="IMAGE:5703909"

/tissue type="whole brain"

/dev stage="embryo 15.5 dpc"

/lab host="DH10B (T1 phage resistant)"

/clone_lib="NIH-BMAP_EW0"

/note="Organ: brain; Vector: pyx-Asc; Site 1: EcoR I;

Site 2: Not I; The library was constructed according to

Bonaldo, Lennon and Soares, Genome Research, 6:791-806,

1996. Denatured mRNA was size fractionated on a 1% agarose

gel. First strand cDNA synthesis was primed with an

oligo-dT primer containing a Not I site. Double stranded

cDNA was size selected according to mRNA size fraction,

ligated with EcoR I adaptor, digested with Not I, and then

cloned directionally into pyx-Asc vector. The library tag

sequence located between the Not I site and the polyA tail

is GTGCGTGGAA. This library was created for the

University of Iowa Mouse Brain Molecular Anatomy Project

(BMAP): 'Gene Discovery in the Developing Mouse Nervous

System', supported by National Institutes of Mental Health

(NIMH), Hemin Chin, Ph.D., program coordinator."

COUNT 167 a 211 c 193 g 139 t 2 others

Y Match 28.6%; Score 574.2; DB 13; Length 712;

Local Similarity 88.6%; Pred. No. 3.4e-115;

hes 632; Conservative 0; Mismatches 80; Indels 1; Gaps 1;

643 CCTTTCAGCATGACAACTGCGACAGCTGTGGAGAGGTGAAGCGGGGTGTTCCAC 702

1 CCTTTGTATGATGACAACTGCGGCGAGTGTGGAGAGGTGAAGCGGGGTGTTCCAC 60

703 ATGCGCATTTATCCCGCCGACTGCCAGAGTCTGTACGGGGCATGATCAGAGTGAC 762

61 ATGCCACATTTATCCCAACAGACTGCCAGAGTCTCTCGTGGCATGATTGAGTGGAT 120

763 GCCGACCGCGCTCACCTAGAGACATTTGAGAAACATATGTTATATAGGGGCAAG 822

121 GCAGCTCGCGCGCTCACCTAGAGACATTTGAGAAACATATGTTATATAGGGGCAAG 180

823 AATGAGCCGGAACAGAGAGCCCATCTCTGCAAGGTGCAAGTCCGCTCGCTGCCAGC 882

181 AATGAGCCGAGAGCCGAGAGCCCATCTCTGCAAGGTGCAAGTCCGCTCGCTGCCAGC 240

883 CTGGAGGACATGACCCCGAGCGTGTGGAGACATGCACTCACTGGGCTGTTCCGAGAC 942

241 TTGGAAGACATTTGACCTGATGTGTGGACAGCATGCACTCACTGGGCTGTTCCGAGAC 300

943 CGCACAGCTGCTCGAGAGCTGTGTGCGAGGAGAGACACCGAGAGATGATTTAC 1002

301 CGCAACAAGCTGCTCGAGAGCTGTGTGCGAGGAGAGATCAGAAAGATGATTTAT 360

1003 TTCTCTCTCTCGAGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1062

361 TTCTCTCTCTCGAGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 420

1063 CGGAACGAGATAGACCTCCCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1122

421 AGGAATGAGATAGACCTCCCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 480

1123 AAGCGCGCCGAG 1182

481 AAGCGCGAGACCTGAGCGCAAGTCCATGGAAGTGTGCTCAGTGTGACAGATGGTGGCTCCCA 540

1183 GTGCTGCGCGCGGCGCATTTGAGATGGCCAGCAGCGGAGAGAGTCTCGGTCCATCAGC 1242

Db 541 GTGCTGCGAG 600

Qy 1243 GTGCTGCTCTCAGGCGCTTTCCACAGCCCACTCAGCAGAGAGAGAGAGAGAGAGAG 1302

Db 601 GTGCTGCTCTCAGGCGCTTTCTACAAGTCCACTCAGCAGTCTCGGCTGAGCCCTCAGCCC 660

Qy 1303 TCACCAAGGGGAGTCCCTCCACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1355

Db 661 TCACCAAGGGGAGTCCCTCCCTTCTACCCCAAGAGAGAGAGAGAGAGAGAGAGAG 712

RESULT 10

BF222739/c

LOCUS

DEFINITION

7q32f02.x1 NCI CGAP GC6 Homo sapiens cDNA clone IMAGE:369987 3'

similar to TR:060843 O60843 PUTATIVE SERINE/THREONINE PROTEIN

KINASE 1, mRNA sequence.

ACCESSION

BF222739

VERSION

BF222739.1 GI:11129830

KEYWORDS

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

ORGANISM

Homo sapiens (human)

REFERENCE

1 (bases 1 to 740)

NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.

AUTHORS

National Cancer Institute, Cancer Genome Anatomy Project (CGAP),

Tumor Gene Index

JOURNAL

Unpublished

COMMENT

Contact: Robert Strausberg, Ph.D.

Email: cgapbs-r@mail.nih.gov

Tissue Procurement: Christopher A. Moskaluk, M.D., Ph.D., Michael

R. Emmert-Buck, M.D., Ph.D.

cDNA Library Preparation: M. Bento Soares, Ph.D., M. Fatima

Bonaldo, Ph.D.

CDNA Library Arrayed by: Greg Lennon, Ph.D.

DNA Sequencing by: Washington University Genome Sequencing Center

Clone distribution: NCI-CGAP clone distribution information can be

found through the I.M.A.G.E. Consortium/LLNL, send email to:

info@image.llnl.gov

High quality sequence stop: 492.

FEATURES

Location/Qualifiers

1..740

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/clone="IMAGE:369987"

/tissue type="pooled germ cell tumors"

/lab host="DH10B"

/clone lib="NCI CGAP GC6"

/note="Vector: pF773D-Pac (Pharmacia) with a modified

polylinker; Site 1: Not I; Site 2: Eco RI; Plasmid DNA

from the normalized library NCI CGAP GC4 was prepared, and

ss circles were made in vitro. Following HAP purification,

this DNA was used as tracer in a subtractive hybridization

reaction. The driver was PCR-amplified cDNAs from a pool

of 5,000 clones made from the same library (cloneIDs

1257096-1258631, 1469064-1470983, and 1475592-1476743).

Subtraction by Bento Soares and M. Fatima Bonaldo.

BASE COUNT 125 a 206 c 241 g 163 t 5 others

ORIGIN

Query Match 28.0%; Score 562.2; DB 10; Length 740;

Best Local Similarity 97.8%; Pred. No. 1.5e-112;

Matches 578; Conservative 0; Mismatches 12; Indels 1; Gaps 1;

Qy 1401 CCCAGAGCTCGGA-GGGGTGCTTGGAGGCGGGCTCAACTCCATCAAGAACAGCTTTC 1459

Db 721 CCCAGAGCTCGAGGGGGTGCCTTGGAGGCGGGCTCAACTCCATCAAGAACAGCTTTC 662

Qy 1460 TGGGCTCACCCCGCTTCCACCGCGGAACTGCAAGTTCGAGCGCGGAGAGATGTCCA 1519

Db 661 TGGGCTCACCCCGCTTCCACCGCGGAACTGCAAGTTCGAGCGCGGAGAGATGTCCA 602

DNA Sequencing by: Agencourt Bioscience Corporation
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:
 http://image.llnl.gov
 Plate: LLAM13496 row: e column: 22
 High quality sequence stop: 681.
 Location/Qualifiers
 1..889
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:6154749"
 /tissue_type="retinoblastoma"
 /lab_host="DH10B (phage-resistant)"
 /clone_lib="NIH_MGC_67"
 /note="Organ: eye; Vector: pCMV-SPORT6; Site 1: NotI;
 Site 2: SalI; Cloned unidirectionally. Primer: Oligo dT.
 Average insert size 1.75 kb. Library constructed by Life
 Technologies."
 207 a 246 c 274 g 160 t 2 others
 Y Match 27.4%; Score 550.6; DB 13; Length 889;
 Local Similarity 82.4%; Pred. No. 5.3e-110;
 Matches 631; Conservative 0; Mismatches 135; Indels 0; Gaps 0;
 59 GCGTGGAGAACGCTGGGCAAGGGGACAGAGTCTGGTGAAGCTGGGGTTCACTGCG 118
 9 GCGTGGAGAACGCTGGGCAAGGGGACAGAGTCTGGTGAAGCTGGGGTTCACTGCG 68
 119 TCACCTGCGAGAGTGGCCATCAGATCGTCAACCGTGAAGAGCTCAGCGAGTGGTGC 178
 69 TCACGGGTGAGAGTGGCCATCAGATCGTCAACCGGAGAGCTGTCGGAGTGGTGC 128
 179 TGATGAAGTGGAGGGGAGATCGGATCTCTGAAGCTCATTTGAGCAGCCCACTGCTTAA 238
 129 TGATGAAGTGGAGGGGAGATCGGATCTCTGAAGCTCATCGAACCCACATGCTCTCA 188
 239 AGCTGACGACGTTTATGAAACAAATAATTTGTACTGTGTGTGCTAGAACACGTTGTGAG 298
 189 AGCTGACGACGTTTATGAAACAAATAATTTGTACTGTGTGTGCTAGAACACGTTGTGAG 248
 299 GTGTGTGAGCTTTGACATCTGTTGAGAGGGGAGGCTGACGCTTAAGGAGGCTCGGA 358
 249 GGGGTGAGCTTTGACATCTGTTGAGAGGGGAGGCTGACGCTTAAGGAGGCTCGGA 308
 359 AGTTCTTCGGGAGATCATCTCTGGGCTGGACTTCTGCCACAGCCACTCCATATGCCACA 418
 309 AGTTCTTCGGGAGATTTGTTCTGGCTGGACTTCTGCCACAGCTACTCCATCTGCCACA 368
 419 GGGATCTGAACCTGAAACCTCTCTGGTGGAGAGAGAGAACAAATTCGCTATCGCAGACT 478
 369 GAGACTTAAGCCCGAGACCTGCTTTTGGATGAGAAAAACAAATTCGCTATCGCAGACT 428
 479 TTGGATGCGTCTCTGACAGTGGCGAGAGCTGTTGAGACAGAGTGTGGTGTCCCCC 538
 429 TCGGATGCGTCTCTGACAGTGGGGAGAGCTTCTTGGAGACAGAGTGTGGTGTCCCCC 488
 539 ACTACGCTTGGCCCGAGGATGATCCGGGGGAGAGATGATGACGCGCGGAGGCGGAGCTGT 598
 489 ATTATGCGTGTCCAGAGGTGATTAAAGGGGAAAAATATGATGCGCCCGGCGAGACATGT 548
 599 GAGCTGCGGCTCATCTCTGCTGCTGCTGGTGGGGCTCTGCTCTGCTGACGATGACA 658
 549 GAGCTGCTGAGTCTATCTCTTCTGCGCTCTCTGCTGGGGCTCTGCTCTGCTGATGACGACA 608
 659 ACTTGGACAGCTGCTGGAGAGAGTGAACGGGGCTGTTTCCACATGCGGACTTTATTC 718
 609 ACTTGGACAGCTGCTGGAGAGAGTGAACGGGGCTGTTTCCACATGCGGCTCTCATTC 668
 719 CCGCCGACTGCCAGAGTCTGCTACGGGGGATGATCGAGGTGAGCGCGGACCGCCGCTCA 778
 669 CTCCAGATTGCCAGAGCTCTCTGAGGGGATGATCGAAGTGGAGCCGAGGAAAAAGGCTCA 728

QY 779 CGTAGAGCACATTGAGAACACATATGGTATATATAGGGGCAAGAA 824
 Db 729 GTCTGGAGCAATTCAGAACATCTTTGTCACNCTAGGCGGGA 774
 RESULT 13
 BQ443644 693 bp mRNA linear EST 29-MAY-2002
 LOCUS UI-M-EW0-bxh-n-01-0-UI.r1 NIH_BMAP_EW0 Mus musculus CDNA clone
 DEFINITION IMAGE:5708352 5', mRNA sequence.
 ACCESSION BQ443644
 VERSION BQ443644.1 GI:21246756
 KEYWORDS ESR.
 SOURCE Mus musculus (house mouse)
 ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 REFERENCE 1 (bases 1 to 693)
 AUTHORS NIH-MGC http://mgc.nci.nih.gov/.
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
 JOURNAL Unpublished
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: rgs@nih.gov
 Tissue Procurement: Dr. James Lin, University of Iowa
 CDNA Library Preparation: Dr. M. Bento Soares, University of Iowa
 CDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
 DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
 Clone Distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:
 http://image.llnl.gov
 This clone was contributed by the Brain Molecular Anatomy Project
 (BMAP)
 Seq primer: pYX-5.
 Location/Qualifiers
 1..693
 /organism="Mus musculus"
 /mol_type="mRNA"
 /strain="C57BL/6"
 /db_xref="taxon:10090"
 /clone="IMAGE:5708352"
 /tissue_type="whole brain"
 /dev_stage="embryo 15.5 dpc"
 /lab_host="DH10B (T1 phage resistant)"
 /clone_lib="NIH_BMAP_EW0"
 /note="Organ: Brain; Vector: pYX-Asc; Site 1: EcoR I;
 Site 2: Not I; The library was constructed according to
 Bonaldo, Lennon and Soares, Genome Research, 6:791-806,
 1996. Denatured mRNA was size fractionated on a 1% agarose
 gel. First strand cDNA synthesis was primed with an
 oligo-dT primer containing a Not I site. Double stranded
 cDNA was size selected according to mRNA size fraction,
 ligated with EcoR I adaptor, digested with Not I, and then
 cloned directionally into pYX-Asc vector. The library tag
 sequence located between the Not I site and the polyA tail
 is GGGCTGGGA. This library was created for the
 University of Iowa Mouse Brain Molecular Anatomy Project
 (BMAP). Gene Discovery in the Developing Mouse Nervous
 System", supported by National Institutes of Mental Health
 (NIMH), Hemin Chin, Ph.D., program coordinator."
 BASE COUNT 168 a 179 c 198 g 147 t 1 others
 ORIGIN
 Query Match 27.3%; Score 548.2; DB 13; Length 693;
 Best Local Similarity 87.8%; Pred. No. 1.7e-109;
 Matches 609; Conservative 0; Mismatches 84; Indels 1; Gaps 1;
 QY 387 GGACTTCTGCGACAGCCACTCCATATGCCAGGATCTGAACCTGAAACCTCTGCT 446
 Db 1 GGACTTCTGCGACAGCCACTCCATATGCCAGGATCTGAACCTGAAACCTCTGCT 60
 QY 447 GGACGAGAGAACCAACATCCGCATCGCAGACTTTGGCATGGCTCCCTGCGAGGTTGGGA 506


```
61 AGATGAGGAGAAC-ACATCGTATTGACAGACTTTGGCATGGCATCCTCGCAGTGGGAGA 119
507 CAGCCTGTTGGAGACGAGCTGGTGGTCCCGCCACTAGCGCTGCCCGAGGTGATCCGGG 566
120 CAGCCTGCTGGAGACGAGCTGGGATCTCCACACTAGCTGCTGCGGAAGTATCCGGG 179
567 GGAGAAAGTATGAGCGCGGAGCGGAGCGGAGCTGGAGCTGGCGGCTCATCTGTTCCGCTT 626
180 CGAAGAGTATGATGGCGCGAGGACAGATGTTGGAGCTGTGGTGTGATCTGTTCCGCTT 239
627 GCTGTGGGGGCTCTGCGCTTCCAGCATGAGCAACTGTTGGACAGCTGCTGGAGAGGTGAA 686
240 GCTGTGGGGGCTCTGCGCTTTCATGATGACAACTGCGGAGCTTCTGGAGAGGTCAA 299
687 GCGGGGCTGTTCCACATGCGGACCTTTATCCCGCGGAGCTGCCAGAGTCTGTACGGG 746
300 GCGTGTGTGTTCCATGATGACATTTATCCACAGACTGCCAGAGTCTCTGCGTGG 359
747 CATGATGAGGTGGAGCGCGCAGCGGCTCAGCTAGAGCAATTCAGAAACACATATG 806
360 CATGATGAGGTGATGAGCTGCGGCGCTCAGCTAGAGCAATTCAGAAACACATATG 419
807 GTATATAGGGGGCAAGATAGCGCCGAGCAAGAGAGCGCCATTCCTGCAAGGTGCAGAT 866
420 GTATATAGGTGGCAAGATGAGCGCGGAGCGGAGCGCCATCCCAAGAGGTGCAGAT 479
867 CCGCTCGCTGCCGAGCTGAGGACATGACCGCGAGCTGCTGGACAGCATGCACTCACT 926
480 .CCGCTCACTACCGAGCTTGGAGACATTTGACCTGTGATGTTGGACAGATGCACTCACT 539
927 GGCTGCTTCCGAGACCGCAACAGCTGCTGAGGAGCTGCTGCTGCGAGGAGGAGAACCA 986
540 GGCTGCTTCCGAGACCGCAACAGCTGCTGAGGAGCTGCTGCTGCTGAGGAGGAGATCA 599
987 GGAGAGATGATTTACTTCTCTCTGAGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1046
600 GGAAGAGATGATTTATTTCTCTCTGATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 659
1047 TGAGGAGCTGCCCGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1080
660 TGAGGAGCTGCCCGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 693
```

BU285204 886 bp mRNA linear EST 27-NOV-2002
603601981F1 CSEQCHN54 Gallus gallus cDNA clone CHEST578a21 5', mRNA
sequence.
BU285204
BU285204.1 GI:25734660
EST.
Gallus gallus (chicken)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Archosauria; Aves; Neognathae; Galliformes; Phasianidae;
Phasianinae; Gallus.
1 (bases 1 to 886)
Boardman, P.E., Sanz-Ezquerro, J., Overton, I.M., Burt, D.W., Bosch, E.,
Pong, W.T., Tickle, C., Brown, W.R.A., Wilson, S.A. and Hubbard, S.J.
A Comprehensive Collection of Chicken cDNAs
Curr. Biol. 12 (22), 1965-1969 (2002)
22335534
12445392
Contact: Simon Hubbard
Department of Biomolecular Sciences
University of Manchester Institute of Science and Technology (UMIST
)
PO Box 88, Manchester, M60 1QD, UK
Tel: 01612008930
Fax: 01612360409
Email: Simon.Hubbard@umist.ac.uk.
Location/Qualifiers
1. .886

RES
source

```
/organism="Gallus gallus"  
/mol_type="mRNA"  
/strain="Compton Line 151"  
/db_xref="taxon:9031"  
/clone="CHEST578a21"  
/sex="Female"  
/tissue_type="not cerebrum or cerebellum"  
/dev_stage="adult"  
/lab_host="DH10B"  
/clone_lib="CSEQCHN54"  
/note="Organ: brain; Vector: pBluescript II KS(+); Site 1:  
EcoRI; Site 2: NotI; This normalized library was  
constructed from 1 million independent clones. cDNA  
synthesis was initiated using an oligo(dT) primer, using  
methylated C in the first strand synthesis reaction.  
Following this first strand reaction, double-stranded cDNA  
was blunted, ligated to NotI adapters, digested with EcoRI  
, size-selected, and cloned into the NotI and EcoRI  
compatible sites of a custom modified MCS of the  
pBluescript (KS+) vector. The library was normalized in 2  
rounds using conditions adapted from Soares et al., PNAS  
(1994) 91: 9228-9232 and Bonaldo et al., Genome Research 6  
(1996): 791, except that a significantly longer  
reannealing hybridization was used."  
BASE COUNT 267 a 174 c 205 t  
ORIGIN  
Query Match 26.0%; Score 522; DB 13; Length 886;  
Best Local Similarity 79.3%; Pred. No. 9.7e-104;  
Matches 657; Conservative 0; Mismatches 165; Indels 6; Gaps 3;  
QY 19 GACGGCGGCGGACGCGGAGTATGTTGGGCGCTACCGGCTGGAGAGACGCTGGGC 78  
Db 2 GGCACGGAGCTCAACATGCAATATGTTGACCCCTATCGTTGGAGAAACCCCTAGA 61  
QY 79 AAGGGCGAGACAGGCTGCTGTAAGCTGGGGGTTCACTGCTCACCTGCCAGAGGTGGCC 138  
Db 62 AAAGGACAGACAGGCGCTTGAAGTTGGGGTTCACTGTGCACATGTCAGAGGTTCG 121  
QY 139 ATCAAGATGCTCAACCGTGAGAGCTCAGCGAGTGGTGTGATGAGTGGAGCGGAG 198  
Db 122 ATAAAAATTTGTCACAGAGAGAGCTCAGCGAGTGGTGTCTAATGAAGGTGGAGCG 181  
QY 199 ATCGGATCCTGAAGCTCATTTGAGCACCCCGCTTAAAGCTGCAGACGCTTATGAA 258  
Db 182 ATAGCCATCTGGAATTTGATAGAGCTCCCATGTTTAAAGCTGCATGACGTTTACG 241  
QY 259 AACAAAAAATTTTGTACTCTGTGTAGAACACGTTGTGAGTGTGAGCTCTTCGACTAC 318  
Db 242 AATAAAAAATTTTGTACTCTGTGTAGAACACGTTGTGAGTGTGAGCTCTTCGACTAT 301  
QY 319 CTGGTGAAGAGGGGAGGCTGACGGCTAAGAGGGCTCGAAGTTCTTCGGGAGATCATC 378  
Db 302 CTTGTAAAAAAGGGAAGATTGACACAAAAGAGCAGGAAGTTCTTCGACAAATCATC 361  
QY 379 TCTGGCTGGAGCTCTGCGACAGCCACTCCATATATGCCACAGGATCTGAAACCTGAAAC 438  
Db 362 TCTGTTTAGACTTCTGCGATAGTCAATATATGCCACAGGACTTAAACCAAGAAAC 421  
QY 439 CTCCTGTGCGAGGAGAAACAAACATCCGCATCGCAGACTTTGGCATGGCGCTCCGAG 498  
Db 422 CTACTGCTGGATGAAAGAACAAATATCCGATAGCAGACTTTGGGATGGCATCGCTGCAG 481  
QY 499 GTTGGGAGAGCTGTTGGAGACGAGCTGTGGTGGTCCCGCCTACGCTGCCCGGAGGTG 558  
Db 482 GTTGGGAGAGCTGTTGAGAAACCAAGTTGTGGATCAGCGCATTTATGCTGCCCTGAA 541  
QY 559 ATCCGGGGGAGAGAGTATGACGCGCGGAGGCGGAGCTGTGGAGCTGGCGGTCACTCTG 618  
Db 542 ATCCGGGGGAGAGAGTATGATGGAAGGAGGAGGAGGAGCTGTGGAGCTGTGGATCAT-TTG 600  
QY 619 TTGCGCTTGTGTTGGGGGCTCTGCGCTTCGACGATGACAACTTCGACAGCTGCTGGAG 678
```

601 GTTCGTTGCTAGTGGTGGTCCCTGCTCCATTCGATGACGACAACTTCCGGCAGCTGTTGGAG 660
679 AAGGTGAACGGGGCGTGTTCACATGCCGCACTTTATCCCGCCGACTGCCAGAGTCTG 738
661 AAGGTGAACGGTGGAGTGTTCACATGCCGCACTTTATCCCTCCGACTGTGAGAACCTC 720
739 CTACGGGCGATGATCGA--GGTGGAGCGCGACGCGCTCAGCTAGACACATTGAGA 796
721 TTGCGGGGATGTGATGAACGTGGGATGCTCGAAACGTTTCACGCTAGAACATATT---A 777
797 AACACATATGTTATAGCGGGCGAAGATGAGCCCGAACCAGAGCAGC 844
778 GAAACCATGGAATATAGGGGTGAAGATTGACGAGAACCAAGAGAAAC 825

15
21
TION 60110773F1 NIH_MGC_16 Homo sapiens cDNA clone IMAGE:3343795 5',
mRNA sequence.
ION BE259121
N BE259121.1 GI:9129619
DS EST.
Homo sapiens (human)
NISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
NCE 1 (bases 1 to 589)
ORS NIH-MGC http://mgc.nci.nih.gov/
NAL National Institutes of Health, Mammalian Gene Collection (MGC)
T Unpublished
Contact: Robert Strausberg, Ph.D.
Email: cgapbs@mail.nih.gov
Tissue Procurement: ATCC
cDNA Library Preparation: Ling Hong/Rubin Laboratory
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at: image.llnl.gov
Plate: LLMC129 row: b column: 20.

ES Location/Qualifiers
1..589
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:3343795"
/tissue_type="retinoblastoma"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH MGC 16"
/note="Organ: eye; Vector: pOTB7; Site 1: XhoI; Site 2:
EcoRI; cDNA made by oligo-dT priming. Directionally
cloned into EcoRI/XhoI sites using the following 5'
adaptor: GGCACGAG(G). Library constructed by Ling Hong
in the laboratory of Gerald M. Rubin (University of
California, Berkeley) using ZAP-cDNA synthesis kit
(Stratagene) and Superscript II RT (Life Technologies).
Note: this is a NIH_MGC Library."
125 a 190 c 177 g 97 t

Y Match 25.7%; Score 515; DB 10; Length 589;
Local Similarity 99.0%; Pred. No. 3e-102;
hes 518; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

1354 CAAAAGGAGAGCCCGCTGGCACGCGCCAAACCCACGCGCCCGTCCAGCCCGCAGCGTCGGA 1413
31 CGAGGGGAGAGCCCGCTGGCACGCGCCAAACCCACGCGCCCGTCCAGCCCGCAGCGTCGGA 90
1414 GGGGTGCTCGTGGAGCGCGGCTCACTCCATCAGACAGCTTTCTGGCTCACCCCGC 1473
91 GGGGTGCTCGTGGAGCGCGGCTCAAGCGCCATCAAGAAACAGCTTTCTGGGTCAACCCCGC 150
1474 TTCCACCGCGGAAACTGCAAGTTCGAGCGCGGAGAGATGTCCAACTGACACACAGAG 1533

Db 151 TTCACCGCCGGAACATGCAAGTTCAGACCCCGAGAGATGTCRAACCTGACACACAGAG 210
Qy 1534 TCGTCCCGCAGAGCTGGCGAAGAAAGTCTCTGTTTGGGAACCTTCATCAGCCTGGAGAGGAG 1593
Db 211 TCGTCCCGCAGAGCTGGCGAAGAAAGTCTCTGTTTGGGAACCTTCATCAGCCTGGAGAGGAG 270
Qy 1594 GAGCAGATCTTCGTGTCATCAAAAGACAAACCTCTGAGCTCCATCAAGGCTGACATCGTG 1653
Db 271 GAGCAGATCTTCGTGTCATCAAAAGACAAACCTCTGAGCTCCATCAAGGCTGACATCGTG 330
Qy 1654 CAGCCCTTCTGTGTCATTCCCACTCTCAGCCACAGCGCTCATCTCCCAAAACGAGCTTCCGG 1713
Db 331 CAGCCCTTCTGTGTCATTCCCACTCTCAGCCACAGCGCTCATCTCCCAAAACGAGCTTCCGG 390
Qy 1714 GCCGAGTACAAGGCCACGGGGGCGCAGCCGTGTTCCAGAAAGCGGTCAGAGTTT 1773
Db 391 GCCGAGTACAAGGCCACGGGGGCGCAGCCGTGTTCCAGAAAGCGGTCAGAGTTT 450
Qy 1774 GATATCACCTACACGGAGGGTGGGAGGCGCAGAGGAGAAACGCGCATCTACTCGTCAAC 1833
Db 451 GATATCACCTACACGGAGGGTGGGAGGCGCAGAGGAGAAACGCGCATCTACTCGTCAAC 510
Qy 1834 TTCACCTGCTCTCAGGCCCCAGCCGTCGCTTCAAGAGGGTGG 1876
Db 511 TTCACCTGCTCTCAGGCCCCAGCCGTCGCTTCAAGAGGGTGG 553

Search completed: November 27, 2003, 11:05:16
Job time : 3816.85 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

cleic - nucleic search, using sw model

n: November 27, 2003, 04:44:37 ; Search time 130.869 Seconds
(without alignments)
6769.052 Million cell updates/sec

ct score: 2007
US-10-054-579-1

nce: 1 atgacatcgacgggaagga.....gaattatccgaaagttaa 2007

ng table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

hed: 569978 seqs, 220691566 residues

number of hits satisfying chosen parameters: 1139956

um DB seq length: 0

um DB seq length: 2000000000

processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

ase : Issued Patents NA.*

1: /cgn2_6/ptodata/2/ina/5A COMB.seq.*

2: /cgn2_6/ptodata/2/ina/5B COMB.seq.*

3: /cgn2_6/ptodata/2/ina/6A COMB.seq.*

4: /cgn2_6/ptodata/2/ina/6B COMB.seq.*

5: /cgn2_6/ptodata/2/ina/BCTUS COMB.seq.*

6: /cgn2_6/ptodata/2/ina/backfiles.seq.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

t	Score	Query Match	Length	DB ID	Description
1	2005.4	99.9	2908	4	US-09-930-181-1
2	1993	99.3	3364	4	US-09-930-181-3
3	225.4	11.2	1742	3	US-08-557-006C-38
4	225.4	11.2	2652	3	US-08-557-006C-39
5	225.4	11.2	2761	3	US-08-557-006C-24
6	213	10.6	213	4	US-09-930-181-5
7	213	10.6	1647	3	US-09-101-146-44
8	212.6	10.6	1675	4	US-09-984-890-1
9	200.4	10.0	1747	3	US-08-557-006C-44
10	191.4	9.5	1929	4	US-09-359-161-4
11	186.4	9.3	2902	4	US-09-579-664B-4
12	176.6	8.8	1736	3	US-08-557-006C-37
13	176.6	8.8	1783	3	US-08-557-006C-36
14	168.2	8.4	2698	2	US-08-677-298-1
15	135.8	6.8	1507	4	US-09-523-849-4
16	135.8	6.8	1544	4	US-09-734-673-1
17	135.8	6.8	2060	4	US-09-523-849-1
18	135.4	6.7	1429	2	US-09-159-385-4
19	135.4	6.7	1429	3	US-09-186-277-4
20	130.8	6.5	2132	2	US-09-159-385-3
21	130.8	6.5	2132	3	US-09-186-277-3
22	121.4	6.0	1619	4	US-09-394-455-14
23	114	5.7	1599	3	US-09-256-465-1
24	114	5.7	1599	4	US-09-167-322-3
25	113.6	5.7	3228	4	US-09-579-664B-5
26	112.4	5.6	2181	4	US-09-417-197-70
27	112.4	5.6	2184	4	US-09-417-197-138

Sequence 1, Appli
Sequence 1, Appli
Sequence 1, Appli
Sequence 1454, Ap
Sequence 526, App
Sequence 12, Appl
Sequence 953, App
Sequence 1, Appli
Sequence 3, Appli
Sequence 15, Appl
Sequence 68, Appl
Sequence 39, Appl
Sequence 1, Appli
Sequence 1, Appli
Sequence 1, Appli
Sequence 1, Appli
Sequence 4, Appli

ALIGNMENTS

RESULT 1
US-09-930-181-1
; Sequence 1, Application US/09930181
; Patent No. 6455292
; GENERAL INFORMATION:
; APPLICANT: Origene Technologies
; TITLE OF INVENTION: Full-Length Serine Protein Kinase in Brain and Pancreas
; FILE REFERENCE: 16U 101 VI
; CURRENT APPLICATION NUMBER: US/09/930,181
; CURRENT FILING DATE: 2001-08-16
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 1
; LENGTH: 2908
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (106)..(2112)
US-09-930-181-1

Query Match 99.9%; Score 2005.4; DB 4; Length 2908;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2006; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
Qy 1 ATGACATCGACGGGGAAGGACGGCGCGGCGAGCAGCGCAGTATGTTGGGCGCTACCGG 60
Db 106 ATGACATCGACGGGGAAGGACGGCGCGGCGAGCAGCGCAGTATGTTGGGCGCTACCGG 165
Qy 61 CTGGAGAAGCGCTGGGCGAAGGGCGGAGAGAGTCTGGTGAAGCTGGGGGTTCACTGGCTC 120
Db 166 CTGGAGAAGCGCTGGGCGAAGGGCGGAGAGAGTCTGGTGAAGCTGGGGGTTCACTGGCTC 225
Qy 121 ACCTGCCAGAGGTGGCCATCAAGATCGTCAACCGGAGAGAGTCTAGCGAGTGGGCTG 180
Db 226 ACCTGCCAGAGGTGGCCATCAAGATCGTCAACCGGAGAGAGTCTAGCGAGTGGGCTG 285
Qy 181 ATGAAGTGGAGCGGAGATCGCGATCCTGAAGCTATTGAGCAGCCCGCCCTCTAAAG 240
Db 286 ATGAAGTGGAGCGGAGATCGCGATCCTGAAGCTATTGAGCAGCCCGCCCTCTAAAG 345
Qy 241 CTGCACAGCTTTATGAAAACAAAATAATTTGTACTGGTGTAGAACACAGTGTAGGT 300
Db 346 CTGCACAGCTTTATGAAAACAAAATAATTTGTACTGGTGTAGAACACAGTGTAGGT 405
Qy 301 GGTGAGCTCTTCGACTACCTGGTGAAGAGGGGAGGCTGACGCGCTAAGGAGGCTCGGAAG 360
Db 406 GGTGAGCTCTTCGACTACCTGGTGAAGAGGGGAGGCTGACGCGCTAAGGAGGCTCGGAAG 465
Qy 361 TTCTTCGGCAGATCATCTCTGCGCTGGAGCTTCTGCCACAGCCACTCCATATGCCACAGG 420

|||||
466 TTCTTCCGCGAGATCATCTCTGCGTGAGCTTTGCGCAAGCCATCTCATATGCCACAGG 525
421 GATCTGAAACCTGAAACCTCTCTGCTGAGCAGAGAAACAATCCGATCCGAGATTT 480
526 GATCTGAAACCTGAAACCTCTCTGCTGAGCAGAGAAACAATCCGATCCGAGATTT 585
481 GGCATGGGCTCTGAGGTTGGGCAACGCTGTTGGAGACAGCTGTTGGTCCCGCCAC 540
586 GGCATGGGCTCTGAGGTTGGGCAACGCTGTTGGAGACAGCTGTTGGTCCCGCCAC 645
541 TACGCTCTCCCGAGGTTGATCCGGGGGAGAGATGATGACGCGGAGGCGGAGCTGTGG 600
646 TACGCTCTCCCGAGGTTGATCCGGGGGAGAGATGATGACGCGGAGGCGGAGCTGTGG 705
601 AGCTCGGGGCTCATCTCTGTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 660
706 AGCTCGGGGCTCATCTCTGTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 765
661 TTGGACAGCTCTGAGAGAGTGAAGCGGGGCTGTTTCCATGCTCCGCACTTTATCCCG 720
766 TTGGACAGCTCTGAGAGAGTGAAGCGGGGCTGTTTCCATGCTCCGCACTTTATCCCG 825
721 CCCGACTGCCAGGTTCTGCTAGCGGCGCATGATCGAGGTTGAGCGCGCACGCGCTCACG 780
826 CCCGACTGCCAGGTTCTGCTAGCGGCGCATGATCGAGGTTGAGCGCGCACGCGCTCACG 885
781 CTAGAGCAATTTAGAGAACATATGTTATATAGGGGGCAAGATGAGCCCGCAACAGAG 840
886 CTAGAGCAATTTAGAGAACATATGTTATATAGGGGGCAAGATGAGCCCGCAACAGAG 945
841 GAGCCATTCCTCGAAGTGCAGATCGCTCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 900
946 GAGCCATTCCTCGAAGTGCAGATCGCTCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1005
901 GAGCTGCTGAGACAGATGACATCACTGCGGCTGCTTCCGAGACGCAACAAGCTGTGCGAG 960
1006 GAGCTGCTGAGACAGATGACATCACTGCGGCTGCTTCCGAGACGCAACAAGCTGTGCGAG 1065
961 GAGCTGCTGCTGAGAGAGAGAACCAAGGAGAGATGATTTACTTCTCTCTGCTGAGCCGG 1020
1066 GAGCTGCTGCTGAGAGAGAGAACCAAGGAGAGATGATTTACTTCTCTCTGCTGAGCCGG 1125
1021 AAGAAAGTACCCGAGCCAGGAGGATGAGGACCTGCCCCCGGCAAGAGATGAGCCCT 1080
1126 AAGAAAGTACCCGAGCCAGGAGGATGAGGACCTGCCCCCGGCAAGAGATGAGCCCT 1185
1081 CCCCAGGAGGCTGTGACTCCCGATGCTGAAACCGGACGCGCAAGCGGCGGCAAGAGCC 1140
1186 CCCCAGGAGGCTGTGACTCCCGATGCTGAAACCGGACGCGCAAGCGGCGGCAAGAGCC 1245
1141 AATCATGAGAGTGTCTAGCTGAGCGGCGGCTCCCGGCTGCTGCTGCTGCTGCTGCTGCT 1200
1246 AAGTCAATGAGAGTGTCTAGCTGAGCGGCGGCTCCCGGCTGCTGCTGCTGCTGCTGCTGCT 1305
1201 ATTGAGTGGCCAGCAGCAGCAGAGGCTCTCGTCCATCAGAGGCTGCTCTCCAGGCTTT 1260
1306 ATTGAGTGGCCAGCAGCAGCAGAGGCTCTCGTCCATCAGAGGCTGCTCTCCAGGCTTT 1365
1261 TCCACAGCCCACTCAGACAGCCCGCGGCTGAGCCCTTCCAGCCCTTCCAGAGGCGGAGTCC 1320
1366 TCCACAGCCCACTCAGACAGCCCGCGGCTGAGCCCTTCCAGCCCTTCCAGAGGCGGAGTCC 1425
1321 CTCGCCACCCCGAGGGGACACTGTTGCTCAACGCGCAAGAGAGCCCGGCTGCGAGCCGC 1380
1426 CTCGCCACCCCGAGGGGACACTGTTGCTCAACGCGCAAGAGAGCCCGGCTGCGAGCCGC 1485
1381 AACCCCAAGCCCGCTCAGCCCGCAGCTGAGGGGCTGCTGAGGGGCGGCTGAGCTCAAC 1440
1486 AACCCCAAGCCCGCTCAGCCCGCAGCTGAGGGGCTGCTGAGGGGCGGCTGAGCTCAAC 1545
1441 TCCATCAAGAACAGCTTTCTGGGCTCACCCCGCTTCCAGCCCGGCAAACTGCAAGTTCCG 1500

Db 1546 TCCATCAAGAACAGCTTTCTGGGCTACCCCGTTCACCGCGGAAACTGCAAGTTCCG 1605
Qy 1501 ACCCGGAGAGATGTCCTCAACCTGACACACAGAGTCTGTCCTCCAGAGTGGCGAAGAGTCC 1560
Db 1606 ACCCGGAGAGATGTCCTCAACCTGACACACAGAGTCTGTCCTCCAGAGTGGCGAAGAGTCC 1665
Qy 1561 TGGTTTGGGAACTTCATCAGCTGAGCTGGAGAGGAGGAGCATCTTCTGTTGTCATCAAGAC 1620
Db 1666 TGGTTTGGGAACTTCATCAGCTGAGCTGGAGAGGAGGAGCATCTTCTGTTGTCATCAAGAC 1725
Qy 1621 AAACCTCTGAGCTCCATCAAGGCTGACATCGTGCACGCTTCTCTGTCGATTTCCAGTCTC 1680
Db 1726 AAACCTCTGAGCTCCATCAAGGCTGACATCGTGCACGCTTCTCTGTCGATTTCCAGTCTC 1785
Qy 1681 AGCACAGGCTCATCTCCAAACGAGCTTCCGGGCGAGATACAAGGCCACGCGGGGGGCCA 1740
Db 1786 AGCACAGGCTCATCTCCAAACGAGCTTCCGGGCGAGATACAAGGCCACGCGGGGGGCCA 1845
Qy 1741 GCGGTGTTCCAGAGCGGTCAGTTTCCAGGTTTCCAGGTTTATACCTACACGAGGCTGGGGAG 1800
Db 1846 GCGGTGTTCCAGAGCGGTCAGTTTCCAGGTTTCCAGGTTTATACCTACACGAGGCTGGGGAG 1905
Qy 1801 GCGCAGAGGAGAACCGCATCTACTCCGTCACTTCCAGCTTCCAGCTTCTCAGGCCCCAGCCGT 1860
Db 1906 GCGCAGAGGAGAACCGCATCTACTCCGTCACTTCCAGCTTCCAGCTTCTCAGGCCCCAGCCGT 1965
Qy 1861 CGTTTCAAGAGGTTGGAGACATCCAGGCCAGCTGCTGAGACACACAGACCCCGCT 1920
Db 1966 CGTTTCAAGAGGTTGGAGACATCCAGGCCAGCTGCTGAGACACACAGACCCCGCT 2025
Qy 1921 GCGGCCAGGACCTTGTGAGACACACCTTAACTGATGAAATGATGACGGGCGGCTTTCC 1980
Db 2026 GCGGCCAGGACCTTGTGAGACACACCTTAACTGATGAAATGATGACGGGCGGCTTTCC 2085
Qy 1981 AAATGTGGAATTTATCCGAAAAGTTAA 2007
Db 2086 AAATGTGGAATTTATCCGAAAAGTTAA 2112

RESULT 2
US-09-930-181-3
; Sequence 3, Application US/09930181
; Patent No. 6455292
; GENERAL INFORMATION:
; APPLICANT: Origene Technologies
; TITLE OF INVENTION: Full-Length Serine Protein Kinase in Brain and Pancreas
; FILE REFERENCE: 16U 101 V1
; CURRENT APPLICATION NUMBER: US/09/930,181
; CURRENT FILING DATE: 2001-08-16
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 3
; LENGTH: 3364
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (482)..(2239)
US-09-930-181-3

Query Match 99.3%; Score 1993; DB 4; Length 3364;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 2007; Conservative 0; Mismatches 0; Indels 4; Gaps 1;
Qy 1 ATGACATCGACGGGAGAGAGCGGCGGCGCAGACGCGCAGTATGTTGGCCCTTACCGG 60
Db 229 ATGACATCGACGGGAGAGAGCGGCGGCGCAGACGCGCAGTATGTTGGCCCTTACCGG 288
Qy 61 CTGGAGAGAGCGCTGGGCAAGGGCGACAGCTGTTGTAAGCTGGGGTTTCACTGCGTC 120
Db 289 CTGGAGAGAGCGCTGGGCAAGGGCGACAGCTGTTGTAAGCTGGGGTTTCACTGCGTC 348
Qy 121 ACCTGCCAGAGAGGTGGCCATCAAGATCGTCAACCGTGAGAGAGCTCAGCGAGTCGTGCTG 180

349 ACC TCC CAG AAG GTGG CCATCA AAG TCGTCA ACGGTGA AAGCTCAG GAGCTCG GTGCTG 408

181 ATGA AGGTGG AGCGGAGATCG CGATCCTGA AGCTCATTTGAG CACACCCACAGTCTCTAAAG 240

409 ATGA AGGTGG AGCGGAGATCG CGATCCTGA AGCTCATTTGAG CACACCCACAGTCTCTAAAG 468

241 CTGCACGAGCTTTATGAAAACAA AATAATTT---GTA CCTGGTGCTAGAACACG TGTC 296

469 CTGCACGAGCTTTATGAAAACAA AATAATTTGTAG GTACCTGGTGTAGAACACG TGTC 528

297 AGGTGGTGA GCTCTTCGACTTACCTGGTGA AAGGGGAGGCTGACG CTTAAGAGAGGCTCG 356

529 AGGTGGTGA GCTCTTCGACTTACCTGGTGA AAGGGGAGGCTGACG CTTAAGAGAGGCTCG 588

357 GAA GTTCTTCCGGCAGATCATCTGCGCTGGA ACTTCCGACAGCAGCACTCCATATGCCA 416

589 GAA GTTCTTCCGGCAGATCATCTGCGCTGGA ACTTCCGACAGCAGCACTCCATATGCCA 648

417 CAGGATCTGAAA CTTGAAAACCTCTGCTGGA CGAAGAAACAA CATCCGATCGCAGA 476

649 CAGGATCTGAAA CTTGAAAACCTCTGCTGGA CGAAGAAACAA CATCCGATCGCAGA 708

477 CTTTGGCATGCGCTCCCTGACAGTTGGCGACAGCTGTTGGAGACACAGCTGTGGTCCCC 536

709 CTTTGGCATGCGCTCCCTGACAGTTGGCGACAGCTGTTGGAGACACAGCTGTGGTCCCC 768

537 CCACTACGCTTCCCGCCGAGGTGATCCGGGGGAGAA GTATGACGCGCGGAAAGCGGACGT 596

769 CCACTACGCTTCCCGCCGAGGTGATCCGGGGGAGAA GTATGACGCGCGGAAAGCGGACGT 828

597 GTGGAGCTGCGGCGTCATCTGTTGCGCTTGTGGTGGGGGCTCTGCGCTTCGACGATGA 656

829 GTGGAGCTGCGGCGTCATCTGTTGCGCTTGTGGTGGGGGCTCTGCGCTTCGACGATGA 888

657 CAACTTGGCAGAGCTGCTGGAAGAGGTGAAGCGGGCGGTTCGACATGCCGCACTTTAT 716

889 CAACTTGGCAGAGCTGCTGGAAGAGGTGAAGCGGGCGGTTCGACATGCCGCACTTTAT 948

717 CCGCGCCGACCTGCCAGAGTCTGCTACGGGGCATGATCGAGGTGGAGCGCGCACCGCGCT 776

949 CCGCGCCGACCTGCCAGAGTCTGCTACGGGGCATGATCGAGGTGGAGCGCGCACCGCGCT 1008

777 CACGCTAGACACATTCAGAAAACACATATGGTATATAGGGGGCAAGATGAGCCCGAAC 836

1009 CACGCTAGACACATTCAGAAAACACATATGGTATATAGGGGGCAAGATGAGCCCGAAC 1068

837 AGAGCAGCCATCTCTGCAAGGTGCAGATCCGCTCGCTGCCAGCCTTGGAGGACATCGA 896

1069 AGAGCAGCCATCTCTGCAAGGTGCAGATCCGCTCGCTGCCAGCCTTGGAGGACATCGA 1128

897 CCGCGAGCTGTGCAGACGATGCACTCACTGGGCTGCTTCCGAGACCGCAACAGCTGCT 956

1129 CCGCGAGCTGTGCAGACGATGCACTCACTGGGCTGCTTCCGAGACCGCAACAGCTGCT 1188

957 GCAGGACCTGCTGTCCGAGGAGGAGAACAGGAGAAAGATGATTTACTTCCTCCTCGGA 1016

1189 GCAGGACCTGCTGTCCGAGGAGGAGAACAGGAGAAAGATGATTTACTTCCTCCTCGGA 1248

1017 CCGGAAAGAAAGGTACCCGAGCCAGAGGATGAGGACCTGCCCCCCCGGAAACGAGATAGA 1076

1249 CCGGAAAGAAAGGTACCCGAGCCAGAGGATGAGGACCTGCCCCCCCGGAAACGAGATAGA 1308

1077 CCTCTCCCGGAAGCGTGTGACATCCCGATGCTGAACCGGCAACGCGAAGCGCGGCCAGA 1136

1309 CCTCTCCCGGAAGCGTGTGACATCCCGATGCTGAACCGGCAACGCGAAGCGCGGCCAGA 1368

1137 AGCAAAATCCATGGAGGTGCTCAGCGTGACGAGCGCGGCTCCCGGCTGCTGCGCGGCG 1196

1369 AGCAAAATCCATGGAGGTGCTCAGCGTGACGAGCGCGGCTCCCGGCTGCTGCGCGGCG 1428

1197 GGCCTATGAGATGGCCAGACAGCGCCAGAGGTTCTGGTCCATCAGCGGTGCTCTCAGG 1256

1429	GGCCATTGAGATGGCCCAAGCAGCGCCAGAGGTTCTCGGTCCATCAGCGGTCCTCTCTCAGG	1488
1257	CTTTTCCACAGGCCACTCAGCAGACCCCGGGTGACCCCTCACCCCTCACCAAGGGGCAG	1316
1489	CTTTTCCACAGGCCACTCAGCAGACCCCGGGTGACCCCTCACCCCTCACCAAGGGGCAG	1548
1317	TCCCTCTCCACCCCAAGGGGACA CTGTGTCACAGGCCAAAGAGAGCCTCGGTGCGAC	1376
1549	TCCCTCTCCACCCCAAGGGGACA CTGTGTCACAGGCCAAAGAGAGCCTCGGTGCGAC	1608
1377	GCCCAACCCACGCCCCCTCCAGGCCCAAGCGCTCGAGGGGTCCCTGAGGGGCGCGCT	1436
1609	GCCCAACCCACGCCCCCTCCAGGCCCAAGCGCTCGAGGGGTCCCTGAGGGGCGCGCT	1668
1437	CAACTCCATCAAGAACAGCTTTCTGGGCTCACCCCGCTTCACCGCCGGAACCTGC AAGT	1496
1669	CAACTCCATCAAGAACAGCTTTCTGGGCTCACCCCGCTTCACCGCCGGAACCTGC AAGT	1728
1497	TCCAGCGCCGAGGAGATGTCCAACTGACACACAGAGTCGTCCCCAGAGCTGGCGAAGAA	1556
1729	TCCAGCGCCGAGGAGATGTCCAACTGACACACAGAGTCGTCCCCAGAGCTGGCGAAGAA	1788
1557	GTCTGTGTTGGGAACTTTCATCAGCCTTGGAGAGGAGGACAGATCTTCGTGGTCATCAA	1616
1789	GTCTGTGTTGGGAACTTTCATCAGCCTTGGAGAGGAGGACAGATCTTCGTGGTCATCAA	1848
1617	AGACAAACCTCTGAGCTCCATCAAGGCTGACATCGTCAGCGCTTCCTGTGATTTCCCGAG	1676
1849	AGACAAACCTCTGAGCTCCATCAAGGCTGACATCGTCAGCGCTTCCTGTGATTTCCCGAG	1908
1677	TCTCAGGCCACAGCGTCATCTCCAAACGAGCTTCCGGGCGGAGTACAAGGCCACGGGGGG	1736
1909	TCTCAGGCCACAGCGTCATCTCCAAACGAGCTTCCGGGCGGAGTACAAGGCCACGGGGGG	1968
1737	GCCAGCGGTGTTCCAGAACGGGTCAAAGTTCAGGTTGATACACCTACACGAGGGGTGG	1796
1969	GCCAGCGGTGTTCCAGAACGGGTCAAAGTTCAGGTTGATACACCTACACGAGGGGTGG	2028
1797	GGAGGCGCAGAGGAAACGGCATCTACTCCGTCACTTCACCGTGTCTCAGCGCCCGAG	1856
2029	GGAGGCGCAGAGGAAACGGCATCTACTCCGTCACTTCACCGTGTCTCAGCGCCCGAG	2088
1857	CCGTCCGCTTCAAGAGGGTGTGAGACCATCCAGGCCACGCTGCTGAGCACAACGACCC	1916
2089	CCGTCCGCTTCAAGAGGGTGTGAGACCATCCAGGCCACGCTGCTGAGCACAACGACCC	2148
1917	GCCTGGGGCCAGCACTTGTTCAGACACCACTTAACCTGTATGGAAATGATACGGGGCGGCT	1976
2149	GCCTGGGGCCAGCACTTGTTCAGACACCACTTAACCTGTATGGAAATGATACGGGGCGGCT	2208
1977	TTCCAAATGTGGAAATTATCCCGAAAAGTTAA	2007
2209	TTCCAAATGTGGAAATTATCCCGAAAAGTTAA	2239

```

RESULT 3
US-08-557-006C-38
; Sequence 38, Application US/08557006C
; Patent No. 6258547
; GENERAL INFORMATION:
; APPLICANT: Beri, Rajjinder K.
; APPLICANT: Carling, David
; APPLICANT: Forder, Robert A.
; TITLE OF INVENTION: NUCLEIC ACID ENCODING AMP-ACTIVATED PROTEIN KINASE
; FILE REFERENCE: NGAP/PHM37598/UST
; CURRENT APPLICATION NUMBER: US/08/557,006C
; CURRENT FILING DATE: 1996-03-06
; PRIOR APPLICATION NUMBER: PCT/GB94/01093
; PRIOR FILING DATE: 1994-05-20
; PRIOR APPLICATION NUMBER: GB 9310489.1
; PRIOR FILING DATE: 1993-05-21
; PRIOR APPLICATION NUMBER: GB 9318010.7
; PRIOR FILING DATE: 1993-08-31

```


BER OF SEQ ID NOS: 44
TWARE: PatentIn Ver. 2.1
ID NO 38
NGTH: 1742
PE: DNA
GANISM: Human AMP protein kinase
557-006C-38

Y Match 11.2%; Score 225.4; DB 3; Length 1742;
Local Similarity 56.5%; Pred. No. 7.6e-42;
hes 440; Conservative 0; Mismatches 336; Indels 3; Gaps 1;
42 GTATGTTGGGCTTACCGCTCGAGAGACGCTGGCAAGGGGCGACAGAGTCTGTGTGAA 101
33 GAAGATCGGACACTAGCTGCTGGGGACACCTTGGCGTGGGACCTTGGGCAAGTGAA 92
102 GCTGGGGTTCCTCGCTCACCCTGCGAGAGGTGGCCATCAAGATCGTCAACCGTGAGAA 161
93 GATTGGAGAAACATCAATTGACAGGCCATTAAGTGGCAGTTAAGATCTTAAATAGACAGAA 152
162 GCTCAGCAGTGGTGTCTGATG---AAGGTGAGCGGGAGATCGGATCCTGAAGCTCAT 218
153 GATTGCGAGTTAGATGTTGTTGGAAAAATTAACAGAGAAATTCAAAAATCTTAAACTCTT 212
219 TGAGCACCCCCAGCTCCTAAAGCTGCAGCAGCTTTATGAAAAACAAAAATATTTGTACT 278
213 TGTCTATCTCATATATCAAACTTACCAAGTATCAGCATCTCCAAACAGACTTTTAT 272
279 GGTGCTAGAACACGTGTCAAGTGGTGTGAGCTTTCGACTACCTGCTGAGAGGGGAGCT 338
273 GGTAAATGGAATATGTTCTGGAGGTGAATGTTTGCAGTACATCTGTAAACACGGGAGGT 332
339 GAGGCTAAGAGGCTCGAAGTCTTCCGCGAGATCATCTCTGCGTGGACTCTTGCCA 398
333 TGAAGAGTGGAGCTCGCGGCTCTTCAGCAGATTTCTGTGCGGTGGACTACTGTGCA 392
399 CAGCCACTCCATATGACACAGGATCTGAAACCTGAAACCTCTGCTGGAGAGAGAA 458
393 CAGGCATGTTGTCCACAGGACTGGAAGCAGAGACGTTGCTGGAGCCGAGAT 452
459 CAACATCCGATCGCAGACTTTGGGATGCGCTCCCTGAGGTGGCGACAGCCTGTGGA 518
453 GAATGCTAAGATAGCTGACTTCGGACTCTTAATATGATGTCAGATGTTGAATTTCTAGC 512
519 GACCACTGTTGGTCCCCCTACAGCTGCGCGGAGTATCGGGGGGAGAGATGA 578
513 AACTAGCTGGATCGCCAAATTAATGACACCGGAGGTCTCTCAGGAAGGCTGTATGC 572
579 CGGCGGAGAGGGGAGCTGTGGAGCTGCGGCTCATCTCTGCTGCTGGTGGGGG 638
573 GGGTCTGAGTTGATATCTGGAGCTGTGGTGTATCTCTGTATGCCCTTCTGTGGCAC 632
639 TCTGCCCTTCGACGATGACAACTTGGCAGCTGTGGAGAGAGGTGAAGCGGGGCTGTT 698
633 CTTCCGGTTCGACGATGAGCAGCTGCTACGCTCTTAAAGAGATCCGAGGGGGTGTGTT 692
699 CCACATGCGGCACTTTATCCCGCCGACTGCCAGATCTGCTACGGGGCATGTCAGGT 758
693 CTACATCCCGGAGTATCTCAACGGTCTATGCGCATCTGCTGATGCAATGCTGCAGT 752
759 GAGCGCGGACGCGGCTCTCAGCTAGACACATTCAGAAACACATATGTTATAGGGG 817
753 GGRACCCCTTGAAGCGAGCACTATCAAGACATACGAGAGCATGATGTTTAAACAGG 811

557-006C-39
ence 39, Application US/0857006C
nt No. 6256547
RAL INFORMATION:
LICANT: Beri, Rajindar K.
LICANT: Carling, David
LICANT: Forder, Robert A.

TITLE OF INVENTION: NUCLEIC ACID ENCODING AMP-ACTIVATED PROTEIN KINASE
FILE REFERENCE: NGAP/PM37588/UST
CURRENT APPLICATION NUMBER: US/08/557,006C
CURRENT FILING DATE: 1996-03-06
PRIOR APPLICATION NUMBER: PCT/GB94/01093
PRIOR FILING DATE: 1994-05-20
PRIOR APPLICATION NUMBER: GB 9310489.1
PRIOR FILING DATE: 1993-05-21
PRIOR APPLICATION NUMBER: GB 9318010.7
PRIOR FILING DATE: 1993-08-31
NUMBER OF SEQ ID NOS: 44
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 39
LENGTH: 2652
TYPE: DNA
ORGANISM: Rat
FEATURE:
NAME/KEY: gene
LOCATION: (1)..(1747)
OTHER INFORMATION: Full length cDNA sequence fragment of Human AMPK -
OTHER INFORMATION: fragment begins at nucleotide 24 and ends with
OTHER INFORMATION: nucleotide 1765
US-08-557-006C-39

Query Match 11.2%; Score 225.4; DB 3; Length 2652;
Best Local Similarity 56.5%; Pred. No. 8.6e-42;
Matches 440; Conservative 0; Mismatches 336; Indels 3; Gaps 1;
QY 42 GTATGTTGGGCTTACCGCTGGAGAGACGCTGGGCAAGGGGCGACAGGTCTGTGTGAA 101
DB 40 GAAGATCGGACACTAGCTGCTGGGGACACCTTGGCGTGGGACCTTGGGCAAGTGAA 99
QY 102 GCTGGGGTTCCTCGCTCACCCTGCGAGAGGTGGCCATCAAGATCGTCAACCGTGAGAA 161
DB 100 GATTGGAGAAACATCAATTGACAGGCCATTAAGTGGCAGTTAAGATCTTAAATAGACAGAA 159
QY 162 GCTCAGCGAGTGGTGTCTGATG---AAGGTGGAGCGGAGATCGGATCCTGGAAGCTCAT 218
DB 160 GATTGCGAGTTAGATGTTGTTGAAAAATTAACAGAGAAATTCAAAATCTTAACTCTT 219
QY 219 TGACACACCCACGCTCTAAAGCTGCACGAGTTTATGAAAAACAAAAATATTTGTACCT 278
DB 220 TCGTCACTCTCATATTAATCAAACTCTACCAAGTGTATCGACTCATCGACACTCCAAACAGACTTTTAT 279
QY 279 GGTCTAGAACACGCTGTGAGTGGTGTGAGCTTTCGACTACTCTGCTGAGAGGGGAGCT 338
DB 280 GGTAAATGGAATATGTTCTGGAGGTGAATGTTTTCGACTACATCTGTAAACACCGGAGGCT 339
QY 339 GACGCTAAGGAGGCTCGGAAGTTCTTCCGCGAGATCATCTCTGCGCTGGACTTCTGCCA 398
DB 340 TGAAGAGGTGGAAGCTCGCGGCTTTCAGAGAGATTCGTCTGCGCTGGACTACTGTCA 399
QY 399 CAGCCACTCCATATGCGACAGAGGATCTGAAACCTGAAACACCTCTGCTGGAACGAGAGAA 458
DB 400 CAGGACATGTTGTTCCACAGGGACCTGAAAGCCAGAGAACGTGTTGCTGGACGCCAGAT 459
QY 459 CAACATCGGATCGCAGACTTGGCATGGGCTCCCTGCGAGTGGCGACAGCCTGTGGA 518
DB 460 GAATGCTAAGATAGTACTGACTTCGGAGTCTCTAATATATGATGTCAGATGTTGAATTTCTACG 519
QY 519 GACCACTGTTGGTCCCCCTACAGCTGCGCGAGGTGATCCGGGGGAGAGATGA 578
DB 520 AACTAGCTGTGATCGCCAAATTAATGACGACCGGAGGTCTCTCAGGAAGGCTGTATGC 579
QY 579 CGGCGGAGAGCGGACGTTGTGAGCTGCGGCTCATCTCTGCTGCCCTTCTGTGGTGGGGG 638
DB 580 GGGTCTGAGGTTGATATCTGGAGCTGTGGTGTATCTCTGTATGCCCTTCTCTGTGGCAC 639
QY 639 TCTGCCCTTCGACGATGACACTTGGACAGCTGCTGAGAGAGGTGAAGCGGGGCTGTT 698
DB 640 CTTCCGTTCCAGCATGAGCACGTGCTCCTTTAAGAGAGATCCGAGGGGGTGTGTT 699
QY 699 CCACATGCGGCACTTTATCCCGCCGACTGCGCAGAGTCTGTACGGGGGATGATCGAGGT 758

700 CTACATCCCGAGTATCTCAACCGTTCTATTGCGCACTCTGCTGATGCAATCTCGAGT 759
 759 GGACGCGGACCGCGCTCAGCTAGAGCAGATTCAGAAACACATATGATATATAGGGG 817
 760 GGACCCCTTGAAGCGGACCACTATCAAGACATACGAGAGCATGAATGTTTAAACAGG 818

557-006C-24
 sence 24, Application US/08557006C
 ent No. 6258547
 ERAL INFORMATION:
 PLICANT: Beri, Rajindar K.
 PLICANT: Carling, David
 PLICANT: Forder, Robert A.
 FILE OF INVENTION: NUCLEIC ACID ENCODING AMP-ACTIVATED PROTEIN KINASE
 JE REFERENCE: NGAP/PHM37588/UST
 RENT APPLICATION NUMBER: US/08/557,006C
 RENT FILING DATE: 1996-03-06
 FOR FILING DATE: 1994-05-20
 FOR APPLICATION NUMBER: PCT/GB94/01093
 FOR APPLICATION NUMBER: GB 9310489.1
 FOR FILING DATE: 1993-05-21
 FOR APPLICATION NUMBER: GB 9318010.7
 FOR FILING DATE: 1993-08-31
 HER OF SEQ ID NOS: 44
 TWARE: PatentIn Ver. 2.1
 ID NO 24
 ENTH: 2761
 (PE: DNA
 GANISM: Artificial Sequence
 NATURE:
 HER INFORMATION: Description of Artificial Sequence: cDNA of rat
 HER INFORMATION: liver AMP protein kinase
 -557-006C-24

Match 11.2%; Score 225.4; DB 3; Length 2761;
 Local Similarity 56.5%; Pred. No. 8.7e-42;
 Matches 440; Conservative 0; Mismatches 336; Indels 3; Gaps 1;
 42 GTATGTTGGCGCTACCGCTGAGAGACGCTGGCAAGGGGACAGAGTCTGTGTGAA 101
 56 GAAGATCGGACACTAGCTGTGGGGACACCTGTGGCGTGGGCGTGGGCAAAAGTGA 115
 102 GCTGGGGTTCATGCTGCTCCTGAGAGTGGCCATCAAGATCGTCAACCGTGAGAA 161
 116 GATTGAGAACATCAATGACAGGCCATAAAGTGGCAGTTAAGATCTTAAATAGACAGAA 175
 162 GCTCAGCGAGTGGTCTGATG---AAGGTGAGCGGAGATCGCGATCCTCAAGCTCAT 218
 176 GATTCGAGTTAGATGTTGTTGGAATAAATAAAGAGAAATTCAAAATCTTAACCTTT 235
 219 TGAGCACCCCGCTCCTAAGCTGACGACGTTTATGAAAAAATAATTTGTACCT 278
 236 TCGTCATCTCATATTAATCAAACTTACCAAGTGTACGACCTCCCAACAGACTTTTTAT 295
 279 GGTGCTAGACACGCTGAGTGGTGGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 338
 296 GGTAAATGGAATATGTTCTGAGAGTGAATTTGTTGCGACTACATCTGTAAACACGGGAGGT 355
 339 GACGCGCTAAGGAGGCTCGGAAGTCTTTCCGGGAGATCATCTCTGCGCTGAGCTTTCTGCCA 398
 356 TGNAGAGTGGAGCTCGCGGCTCTTCAGCAGATCTGCTGCGCTGAGTACTGTCA 415
 399 CAGCCCATCTCATATGACAGGATCTGAAACCTGAAACCTCTGCTGCTGAGCAGAGAA 458
 416 CAGGCACATGTTGTCCACAGGACCTGAAGCCAGAGAACGTTGCTGAGCGCCAGAT 475
 459 CACATCCGATCGACAGCTTTGGCATGGGCTCCCTGCGAGTTGGCGACAGCTGTTGA 518
 476 GAATGCTAAGATAGCTGACTTCGGAATCTCTTAATATGATGTCAGATGGTGAATTTTACG 535

QY 519 GACAGCTGTGGTCCCGCCACTAGCTGCTCCCGAGTATCCGGGGGAGAAAGTATGA 578
 DB 536 AACTAGCTGTGGTCCCGCCAAATATGAGCAACCGAGGTCACTCAGGAAGCTGTATGC 595
 QY 579 CGGCGGGAAGCGGACGCTGTGGAGCTGCGGCGTCACTCTGTTTCGCTTGTGGGGGC 638
 DB 596 GGTCTCTGAGGTTGATATCTGAGCTGTGGTGTATCTCTGTTATGCGCTTCTCTGTGGCAC 655
 QY 639 TCTGCCCTTCGACGATGACAACTTGGCAGACGCTCTGAGAGAGGTGAAGCGGGCGTCTT 698
 DB 656 CTTCCCGTTTCAGCATGAGCAGCTGCTTCTTTAAGAGATCCGAGGGGGTGTCTT 715
 QY 699 CCACATGCCGCACTTTATCCCGCCGACTGCGCAGAGTCTGCTACGGGCGCATGATCGAGT 758
 DB 716 CTACATCCGAGTATCTCAACCGTTCTATTGCCACCTGCTGATGACATGCTGCAGT 775
 QY 759 GGACGCGCACCGCCCTCAGCTAGAGCAATTCAGAAAAACATATATATATAGGGG 817
 DB 776 GGACCCCTTGAAGCGGACCACTATCAAAAGACATACGAGCATGAATGTTTAAACAGG 834

RESULT 6
 US-09-930-181-5
 ; Sequence 5, Application US/09930181
 ; Patent No. 6455292
 ; GENERAL INFORMATION:
 ; APPLICANT: Origene Technologies
 ; TITLE OF INVENTION: Full-length Serine Protein Kinase in Brain and Pancreas
 ; FILE REFERENCE: 160 101 V1
 ; CURRENT APPLICATION NUMBER: US/09/930.181
 ; CURRENT FILING DATE: 2001-08-16
 ; NUMBER OF SEQ ID NOS: 18
 ; SOFTWARE: PatentIn version 3.0
 ; SEQ ID NO 5
 ; LENGTH: 213
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 ; FEATURE:
 ; NAME/KEY: CDS
 ; LOCATION: (1)..(213)
 US-09-930-181-5

Query Match 10.6%; Score 213; DB 4; Length 213;
 Best Local Similarity 100.0%; Pred. No. 2.6e-39;
 Matches 213; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 ATGACATCGACGGGGAAGGACGCGCGCGCAGCAGCAGTATGTTGGGCGCTTACCGG 60
 DB 1 ATGACATCGACGGGGAAGGACGCGCGCGCAGCAGCAGTATGTTGGGCGCTTACCGG 60
 QY 61 CTGGAGAGACGCTGGGCAAGGGGCGAGAGTCTGTTGAAGCTGGGGGTTCACTGCGTC 120
 DB 61 CTGGAGAGACGCTGGGCAAGGGGCGAGAGTCTGTTGAAGCTGGGGGTTCACTGCGTC 120
 QY 121 ACCTGCCAAGAGTGGCCATCAAGATCGTCAACCGTGAGAGCTCAGCGAGTCCGCTGTG 180
 DB 121 ACCTGCCAAGAGTGGCCATCAAGATCGTCAACCGTGAGAGCTCAGCGAGTCCGCTGTG 180
 QY 181 ATGAAGGTGGAGCGGAGATCCGATCTCTGAAG 213
 DB 181 ATGAAGGTGGAGCGGAGATCCGATCTCTGAAG 213

RESULT 7
 US-09-101-146-44
 ; Sequence 44, Application US/09101146
 ; Patent No. 6124125
 ; GENERAL INFORMATION:
 ; APPLICANT: Dartmouth College, St. Vincent's Institute of
 ; APPLICANT: Medical Research, Kemp et al.
 ; TITLE OF INVENTION: No. 6124125el AMP Activated Protein Kinase
 ; NUMBER OF SEQUENCES: 64
 ; CORRESPONDENCE ADDRESS:

ADDRESSEE: Jane Massey Licata, Esq.
STREET: 66 E. Main Street
CITY: Marlton
STATE: NJ
COUNTRY: USA
ZIP: 08053

COMPUTER READABLE FORM:

MEDIUM TYPE: DISKETTE, 3.5 INCH, 1.44 MB STORAGE
COMPUTER: IBM PC

OPERATING SYSTEM: WINDOWS 95

SOFTWARE: WORDPERFECT 6.0 FOR WINDOWS

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/101.146

FILING DATE: October 7, 1998

CLASSIFICATION: 435

RICOR APPLICATION DATA:

APPLICATION NUMBER: PN7450

FILING DATE: 8 JAN 1996

ATTORNEY/AGENT INFORMATION:

NAME: Jane Massey Licata

REGISTRATION NUMBER: 32,257

REFERENCE/DOCKET NUMBER: DC-0050

ELECOMMUNICATION INFORMATION:

TELEPHONE: (856) 810-1515

TELEFAX: (856) 810-1454

ORAMATION FOR SEQ ID NO: 44:

SEQUENCE CHARACTERISTICS:

LENGTH: 1647

TYPE: Nucleic acid

STRANDEDNESS: Single

TOPOLOGY: Linear

NTI-SENSE: No

101-146-44

Y Match 10.6%; Score 213; DB 3; Length 1647;
Local Similarity 55.6%; Pred. No. 4.7e-39;
hes 430; Conservative 0; Mismatches 340; Indels 3; Gaps 1;

42 GTATGTTGGGCTTACCGCTGGAGAACGCTGGCAAGGGGAGAGAGTCTGTGGTAA 101

33 GAAGATCGGCCACTACATCTCTGGGGGACACGCTGGCGCTCGGCACCTTCGGGAAAGTAA 92

102 GCTGGGGTTCACCTGCTGCTGATGAAGGTGGAGGGGAGATCGCATCTTGAAGCTCAT 161

93 GGTGGGCAAGCAGAGTTGACTGGACATAAAGTTCTGTGAAGATCTCAACCGGAGAA 152

162 GCTC---AGCGAGTGGTCTGATGAAGGTGGAGGGGAGATCGCATCTTGAAGCTCAT 218

153 GATTGGAAGCTGGAGCTGCTGGGAAATCCGACAGAGATCCAGAACCTTGAAGCTTTT 212

219 TGAGCACCCACGCTCTAAAGCTGCACAGCTTTATGAAACAAAAATATTTGTACCT 278

213 CAGGCACCTCATATAATCAAACTGTACAGGTCTATGATACACCGTCTGTATTTTCAT 272

279 GGTGCTAGAACACGCTGTCAGGTGGTGTGAGCTCTTGCATACCTGCTGGAAGGGAGCT 338

273 GGTATGGAATATGCTCAGGAGGAGCTATTTGATATATCTGTAAATATGGAAGTT 332

339 GACGCTAAGGAGGCTCGGAAGTTCTTCGGGAGATCATCTCTGCGCTGGACTTCTTGCCA 398

333 GGAAGAAAGGAGAGTTCGAGCTGCTGTCAGCAGATCTTTCTGTGTGGACTATTGTCA 392

399 CAGGCATCTCATATGACAGGAGTCTGAAACCTGAAACCTCTGCTGGAGAGAGAA 458

393 CAGGCATATGTTGGTCTCAGAGAGATTGAAACCTGAAACCTGCTGTGTATGACACAT 452

459 CAACATCCGATCGCAGACTTTGGATGTCGCTGCTGAGGTGGCGACAGCTTGTGGA 518

453 GAATGCAAAAGATGAGCGAGCTTCGCTTTCAAAATGATGATGATGATGATTTTAA 512

519 GACCAAGCTGTGGTCCCCCACTACGCTCTGCCCGAGGTGATCCGGGGGAGAGATGA 578

513 AACGAGCTGTGCTCGCCCAATTTATGCTGCACCAAGAGTAAATTTTCAGGAAGATTCTACGC 572

QY 579 CGGCGGAAAGCGAGCGTGTGAGCTGGCGCTCATCTCTGTTCCGCTTCTGTTGGGGC 638
DB 573 AGGCCCTGAAGTAGACATCTGGAGCAGCGGGTCACTCTATGCTTCTGTTGGAAAC 632
QY 639 TCTGCCCTTCGACGATGACAACTTGGACAGCTGCTGGAGAAGGTGAACGGGGCTGTT 698
DB 633 TCTCCCTTTTGTATGATGACACGTCGCAACTCTTTTAAAGATATGTGACGGATATT 692
QY 699 CCACATGCGCACCTTTATCCGCCGACCTGCCAGAGTCTGTACGGGGCATGATCGAGT 758
DB 693 TTATACCCCTCAGTATTTGAATCCCTCTCTGTAATAGCTTTTGAACATATGTCAGGT 752
QY 759 GGACGGCGGACCGGCTCAGCTAGAGCACATTCAGAAACACATATGTTATA 811
DB 753 AGATCTATGAAGAGGGCCCAATAAAGATATCAGGGAACATGATGTTTA 805

RESULT 8

US-09-984-890-1

; Sequence 1, Application US/09984890

; Patent No. 6492156

; GENERAL INFORMATION:

; APPLICANT: YAN, Chunhua et al.

; TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC

; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES

; TITLE OF INVENTION: THEREOF

; FILE REFERENCE: CL001306

; CURRENT APPLICATION NUMBER: US/09/984,890

; CURRENT FILING DATE: 2001-10-31

; NUMBER OF SEQ ID NOS: 4

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 1

; LENGTH: 2175

; TYPE: DNA

; ORGANISM: Homo sapiens

US-09-984-890-1

Query Match 10.6%; Score 212.6; DB 4; Length 2175;
Best Local Similarity 54.2%; Pred. No. 6.3e-39;
Matches 431; Conservative 0; Mismatches 364; Indels 0; Gaps 0;

QY 47 TTGGGCGCTTACCGCTGGAGAACGCTGGGCAAGGGGAGAGCTCTGTGTTGAAGCTGG 106

DB 149 TTGGAATCTACCGCTCTCTCAAGACCAATTTGGCAAGGTAAATTTTGGCAAGGTGAAGTTGG 208

QY 107 GGGTTTCACTGCTCAGCTGCTGCAAGAGGTGGCCATCAAGATCTGCAACCGGTGAGAAGCTCA 166

DB 209 CCCGACACATCTGACTGGGAAAGAGGTAGTCTGTGAAGATCAATTGACAGACTCAACTGA 268

QY 167 GCGAGTCGCTGATGAAGGTGGAGGGGAGATCGGATCTCTGAAGCTCATTTGAGCACC 226

DB 269 ACTCTCCAGCTCCAGAAACTATTCCGGAAGTAAAGATAATGAAAGGTTTGAATCATC 328

QY 227 CCCAGCTCTTAAAGCTGACGAGCTTTATGAAACAAAAATATTTGTACTCTGTGCTAG 286

DB 329 CCAACATAGTTAAATTTTGAAGTATTTGAGACTGAGAAAACGCTCTACTCTTGTCTAGG 388

QY 287 AACACGCTGTCAGGTGGTGTGCTCTTCTGACTTACCTGTGTAAGAGGGGAGGCTGACGCTTA 346

DB 389 AGTACGCTAGTGGCGGAGAGTATTTGATTTACCTTAGTGGCTCATGGCAGGATGAAAGAAA 448

QY 347 AGGAGGCTCGGAAGTCTTCCGGCAGATCATCTCTGGCTGGACTTCTGCCACAGCCACT 406

DB 449 AAGAGGCTCGAGCCAAATTCGCCCAGGTAGTGTCTGTGTGAGTACTGTCTCAGCAGAAAT 508

QY 407 CCATATCCACAGGAGTCTGAAACCTGAAACCTCTCTGCTGGAGAGAGAAACAAACATCC 466

DB 509 TTATGTGCTATAGAGACTTAAAGCAGAAACCTGCTCTTGGATGCTGATATGACATCA 568

QY 467 GCATCGCAGACTTTGGCATGGCGTCCCTCGAGGTTGGCGACAGCTGTTGGAGACCAGCT 526

DB 569 AGATTGAGACTTTGGCTTCAGCAAATGAAATTCACCTTTTGGAAACAAAGCTGGAACACCTTCT 628

527 GTGGTCCCTCCACCTACGCTGCCCCGAGGTGATCCGGGGGAGAGATGATGACGCGCCGA 586
 629 GTGGCAGTCCCTTATGCTGCTCCCGCAGAACTCTTCCAGGGCAAAAAATATGATGGACCG 688
 587 AGGCGGAGTGTGGAGCTGCGGCTGATCTCTGTTGCGCTTCTGTTGGGGGCTCTGGCCT 646
 689 AGGTGGATGTGTGGAGCTTACGAGTTATCTCTATACACTGGTCAGCGGATCCCTGGCCTT 748
 647 TCGAGATGACAACTTTCGACAGCTGCTGGGAGAGGTGAAGCGGGCGTGTTCACATGC 706
 749 TTGATGGACAGAACTCAAGAGCTTCCGGAAACGGTACTGAGGGGAAAAATACCGTATTC 808
 707 CGCACTTTATCCCGCCGAGCTGCCAGAGTCTGTACGGGGGATGATCAGGTGGAGCGCG 766
 809 CATTCTACATGTCCAGCGACTGTGAACACCTGCTTAAAGAAATTTCTCATTTAATCCA 868
 767 CACGCGCTCAGCTAGACGACATTCAGAAACACATATGTTATATAGGGGGCAGAAATG 826
 869 GCAAGAGAGGCACTTTAGAGCAAAATCATGAAGATCGATGGATGAATGTGGGTCAAG 928
 827 AGCCCGAACCAAGC 841
 929 ATGATGAATAAGC 943

AT 9
 I-557-006C-44
 rrence 44, Application US/08557006C
 ent No. 6258547
 ERAL INFORMATION:
 PLICANT: Beri, Rajindar K.
 PLICANT: Carling, David
 PLICANT: Forder, Robert A.
 TLE OF INVENTION: NUCLEIC ACID ENCODING AMP-ACTIVATED PROTEIN KINASE
 LE REFERENCE: NCAP/PH37588/UST
 RRENT APPLICATION NUMBER: US/08/557,006C
 RRENT FILING DATE: 1996-03-06
 RIOR APPLICATION NUMBER: PCT/GB94/01093
 RIOR FILING DATE: 1994-05-20
 RIOR APPLICATION NUMBER: GB 9310489.1
 RIOR FILING DATE: 1993-05-21
 RIOR APPLICATION NUMBER: GB 9318010.7
 RIOR FILING DATE: 1993-08-31
 MBER OF SEQ ID NOS: 44
 FTWARE: PatentIn Ver. 2.1
) ID NO 44
 LENGTH: 1747
 TYPE: DNA
 ORGANISM: Human
 FEATURE:
 NAME/KEY: gene
 OCATION: (1)..(1747)
 YHER INFORMATION: Full length cDNA sequence fragment of Human AMPK -
 YHER INFORMATION: fragment begins at nucleotide 24 and ends with
 YHER INFORMATION: nucleotide 1765
 I-557-006C-44

try Match 10.0%; Score 200.4; DB 3; Length 1747;
 Local Similarity 56.1%; Pred. No. 3.3e-36;
 ches 440; Conservative 0; Mismatches 336; Indels 8; Gaps 3;
 42 GTATGTTGGGCGCTACCGGCTGGAGAACGCTGGGCAAGGGGAGACAGGCTCTGGTGAA 101
 33 GAAGATCGACACTACGCTGCTGGGGACACCTTGGCGCTCGGCACCTTCGCAAGTGAA 92
 102 GCTGGGGTTCATCGGCTCACTGTCAGAGAGGTGGCCATCAAGATGTCACCGTGAGAA 161
 93 GATTGGAGAACATCAATTGACAGGGCCATAAAGTGGCAGTTAAGATCTTAAATAGACAGAA 152
 162 GCTCAGCGAGTGGTGGCTGATG---AAGTGGAGGGGAGATCGGATCTCTGAGCTCAT 218
 153 GATTGCGAGTTTAGATGTTGTTGGAAAAATAAAACGAGAAATTCAAAATCTTAAACTCTT 212

QY 219 TGAGCACCCCAAGTCTCTAAAGCTGACAGCGTTTATGAAAAACAAAAATATTGTACCT 278
 DB 213 TCGTCATCTCATATATCAAACTTACCAAGTATCAGCACTCCACAGACTTTTAT 272
 QY 279 GGTGCTAAGAACAGCTGTGAGTGTGAGCTCTTCGACTACTCTGGTGAAGAGGGAGGCT 338
 DB 273 GGTAAATGGAATATGTGTGAGGTGAATTTGTTGACTACATCTGTAAACACAGGGAGGT 332
 QY 339 GACGCTTAAGAGGCTCGGAAGTCTTCGCGGAGATCATCTCTGCGCTGGACTTCG--C 396
 DB 333 TGAAGAGGTGAAAGCTCGCGGCTCTTCCAGCAGATTCTGTCTGCCGTGGACTACTGTCT 392
 QY 397 CACAGCCACTCCATATATGCCACAGGGATCTGAAACCTGAAACCTCTCTGCTGGAGAGAG 456
 DB 393 CACAGGCACATGGTTGTCCACAGGGACCTGACGAGAGAACTGTTGCTGGAGCCGAG 452
 QY 457 AACAAATCCGCATCGCAGACTTTGGGATCGGCTCTGAGGTTCGCGACAGCCTTTG 516
 DB 453 ATGAAATGCTAAGATAGCTGACTTCGGACTCTCTAAATATGATGTCAATGTTGAATTTCTA 512
 QY 517 GAGACCAAGCTGTGGTCCCCCTCCACTAGCTGCGCCGAGGTGATCCGGGGGAGAGATAT 576
 DB 513 CGAACTAGCTGTGATCGCCAAATTTATGACAGCCGAGGTCTCTCAGGAAGCTGTAT 572
 QY 577 GACGGCGGAGAGCGGACGCTGTGGAGCTGCGGCTCATCTCTGCTTGCCTTGCCTGCTGGG 636
 DB 573 GCGGCTCTGAGGTTGATATCTGGAGCTGTGGTGTATCTGTATGCTTCTCTGTGGC 632
 QY 637 GCTTGCCTTT---CGACGATGACAACTTCGACAGCTGTGGAGAGGTGAAGCGGGG 693
 DB 633 ACCTCCGTTTCGACGAGATGAGACGCTGCTCTTTAAGAGATCCGAGGGGT 692
 QY 694 GTGTTCCACATCGCGACTTTATCCCGCCGAGTCCGAGAGTCTGCTACCGGGCATATC 753
 DB 693 GTGTTTACATCCCGAGTATCTCAACCGTTCTATTGCCACTCTGTGATGCACATCTG 752
 QY 754 GAGGTGACCGCCGACGCGCTCAGCTAGAGACATTCAGAAACACATATGTTATATA 813
 DB 753 CAGGTGACCCCTTGAAGCGAGCACTATCAAGACATACAGAGCATGATGGTTAAA 812
 QY 814 GGGG 817
 DB 813 CAGG 816

RESULT 10
 US-09-359-161-4
 ; Sequence 4, Application US/09359161A
 ; Patent No. 6342656
 ; GENERAL INFORMATION:
 ; APPLICANT: Bradford, Kent J.
 ; APPLICANT: Dahal, Peetambar
 ; APPLICANT: Yang, Hong
 ; APPLICANT: Cooley, Michael
 ; APPLICANT: Downie, Bruce
 ; APPLICANT: Gee, Oliver
 ; APPLICANT: The Regents of the University of California
 ; TITLE OF INVENTION: Regulation of Source-Sink Relationships and Responses
 ; TITLE OF INVENTION: to Stress Conditions in Plants
 ; FILE REFERENCE: 023070-095900US
 ; CURRENT APPLICATION NUMBER: US/09/359,161A
 ; NUMBER OF SEQ ID NOS: 7
 ; SOFTWARE: PatentIn Ver. 2.1
 ; SEQ ID NO 4
 ; LENGTH: 1929
 ; TYPE: DNA
 ; ORGANISM: Lycopersicon esculentum
 ; FEATURE:
 ; OTHER INFORMATION: Lycopersicon esculentum plant homolog of yeast
 ; OTHER INFORMATION: SNF1 kinase subunit of protein kinase (LeSNF1)
 US-09-359-161-4

1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40	41	42	43	44	45	46	47	48	49	50	51	52	53	54	55	56	57	58	59	60	61	62	63	64	65	66	67	68	69	70	71	72	73	74	75	76	77	78	79	80	81	82	83	84	85	86	87	88	89	90	91	92	93	94	95	96	97	98	99	100
1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40	41	42	43	44	45	46	47	48	49	50	51	52	53	54	55	56	57	58	59	60	61	62	63	64	65	66	67	68	69	70	71	72	73	74	75	76	77	78	79	80	81	82	83	84	85	86	87	88	89	90	91	92	93	94	95	96	97	98	99	100
1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40	41	42	43	44	45	46	47	48	49	50	51	52	53	54	55	56	57	58	59	60	61	62	63	64	65	66	67	68	69	70	71	72	73	74	75	76	77	78	79	80	81	82	83	84	85	86	87	88	89	90	91	92	93	94	95	96	97	98	99	100
1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40	41	42	43	44	45	46	47	48	49	50	51	52	53	54	55	56	57	58	59	60	61	62	63	64	65	66	67	68	69	70	71	72	73	74	75	76	77	78	79	80	81	82	83	84	85	86	87	88	89	90	91	92	93	94	95	96	97	98	99	100
1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40	41	42	43	44	45	46	47	48	49	50	51	52	53	54	55	56	57	58	59	60	61	62	63	64	65	66	67	68	69	70	71	72	73	74	75	76	77	78	79	80	81	82	83	84	85	86	87	88	89	90	91	92	93	94	95	96	97	98	99	100
1	2	3	4	5	6	7	8	9	10	11	1																																																																																								

P 11
-579-664B-4
rence 4, Application US/09579664B
ent No. 6514719
ERAL INFORMATION:
PLICANT: Immunex Corporation
PLICANT: Bird, Timothy A.
PLICANT: Virca, G. Duke
PLICANT: Martin, Urja
PLICANT: Anderson, Dirk M.
FILE OF INVENTION: NOVEL MURINE AND HUMAN KINASES
RE REFERENCE: 2923-A
CURRENT APPLICATION NUMBER: US/09/579,664B
CURRENT FILING DATE: 2000-05-26
NUMBER OF SEQ ID NOS: 36

```

; SOFTWARE: PatentIn version 3.1.1
; SEQ ID NO 4
; LENGTH: 2902
; TYPE: DNA
; ORGANISM: Mus musculus
US-09-579-664B-4

```

Query Match 9.3%; Score 186.4; DB 4; Length 2902;
Best Local Similarity 53.5%; Pred. No. 5.5e-33;
Matches 413; Conservative 0; Mismatches 356; Indels 3; Gaps 1;

Qy	52	CCCTACCGCGTGGAGAAGACGCTGGCGCAAGGGCGACACAGTCTCTGTGAAGCTTGGGGTT	111
Db	288	CGCTTACGAGTTCCTGGAGACGCTGGCGAAGGGCACTTACGGGAAGGTGAAGAAGGCACGA	347
Qy	112	CACTGGGTCACTGCCAGAAAGTGGCCATCAAGATCGTCAACCGTCGAGAAGCTCAGCGAG	171
Db	348	GAGAGCTCGGGGGCTCTGGTGGCCATCAAGTCATCAGGAAGAACAATCAAGAATGAG	407
Qy	172	TGCGGTGCTGATGAAGGTGGAGCGGGAGATCGCGATCCTGAAGCTCATTTGAGCAACCCAC	231
Db	408	CAGGATCTGTCACATACGGAGGGAGATTGAGATCATGTCTTCACTCAACCAACCCCCAC	467
Qy	232	GTCTTAAAGTGCAGACGTTTATGAAAAAATAATTTGTACTCTGTGCTGTAGAAACAC	291
Db	468	ATCATTTGCCATCCTCAAGTGTTTGAGAAATAGCAGCAAGATTGTGATTTGTCTATGGAGTAT	527
Qy	292	GTGTCAAGTGGTGAGTCTTCGACTACCTGGTCAAGAAAGGGGAGGCTGACGCCTAAGGAG	351
Db	528	GCAGCGGAGGGGATCTGTATGATTACATCAGTGAGCGGCACCGCTGAGTGAGCGGGAC	587
Qy	352	GCTCGGAAGTTCTTCCGGCAGATCACTCTCGCGCTGGACTTCTGCCACAGCCACTCCATA	411
Db	588	GCCAGGCAATTTCTTCCGACAGATCGTGTCTGCCCTGCACTACTGCCACCAGACGGGATC	647
Qy	412	TGCCACAGGGATCTGAAACCTGAAAAACCTCTCTGCTGGAAGAAAGAACCAATCCCGATC	471
Db	648	GTTTCACCGAGATCTCAAGCTGGAAAAACATCTTCTAGATGCCCAATGAAAAATCAAGATT	707
Qy	472	GCAGACTTTGGCATGGGCTCCGTCAGAGTTGGCGACAGCCTGTTTGAGACACAGCTGTGGG	531
Db	708	GCTGACTTTGGCCTCTCCAACTGTACCAAAAGGCCAAGTTCTCTCAGACGTTCTGTGGG	767
Qy	532	TCCCGCCCACTACGCTGCCCGAGGTGATCCGGGGGGAGAAATGATACGCCGCCGGAAGCG	591
Db	768	AGCCCTCTCTACGCTCGCTGAGATAGTCAACGGGNAAGCCCTATGTGGGCCCAGAGGTG	827
Qy	592	GACGTGTGAGCTGCGCGGTGATCTCTTTCGCTTTCGCTGTGGGGGCTGTGCCCTTCGAC	651
Db	828	GACAGCTGGTCTCTGGGCGTTCTCTGTATACCTCTGTGATGGCACCATGCTCTTTGAC	887
Qy	652	GATGACAACTTGGGACAGCTGTGGAGAGGTGAAGCGGGCGTGTTCACATGCCCGAC	711
Db	888	GGGCAGGATCATAAAAACACTGGTGAAGCAAAATCAGTAACGGGGCTTACCGTGAAGCGCCC	947
Qy	712	TTTATCCCGCCGACTGCCAGAGTCTGCTACGGGGCATGATCGAGGTGGAACGCCGACGC	771
Db	948	---AAGCGTCCGATGCTGTGCGCTGATCCGCTGGCTGTTAATGTGTGAACCCACCCGT	1004
Qy	772	CGCCTCACGCTAGACACATTCAGAAACACATATGTGTATATAGGGGGCAAGA	823
Db	1005	CGGGCCACACTGGAGATGTAGCAGTCAATGGTGGGTCACTGGGGGTACA	1056

RESULT 12
US-08-557-006C-37
Sequence 37, Application US/08557006C
Patent No. 6259547
GENERAL INFORMATION:
APPLICANT: Beri, Rajendar K.
APPLICANT: Carling, David
APPLICANT: Fordner, Robert A.
TITLE OF INVENTION: NUCLEIC ACID EN

LE REFERENCE: NGAP/PHM37588/UST
RENT APPLICATION NUMBER: US/08/557,006C
RENT FILING DATE: 1996-03-06
IOR APPLICATION NUMBER: PCT/GB94/01093
IOR FILING DATE: 1994-05-20
IOR APPLICATION NUMBER: GB 9310489.1
IOR FILING DATE: 1993-05-21
IOR APPLICATION NUMBER: GB 9318010.7
IOR FILING DATE: 1993-08-31
MBER OF SEQ ID NOS: 44
FTWARE: PatentIn Ver. 2.1
ID NO 37
LENGTH: 1736
YPE: DNA
GANISM: Artificial Sequence
ATURE:
THER INFORMATION: Description of Artificial Sequence: Rat liver AMP
THER INFORMATION: protein kinase
-557-006C-37

fy Match 8.8%; Score 176.6; DB 3; Length 1736;
t Local Similarity 52.6%; Pred. No. 7.8e-31;
ches 408; Conservative 0; Mismatches 364; Indels 3; Gaps 1;
42 GTATGTTGGCCCTACCGCTGGAGAGACGCTGGGCAAGGGCGACAGAGCTCTGTGAA 101
33 GAAGATCGGACACTAGCTGCTGGGCGACACGCTGGGCGTGGCACCTTCGGCAAGTGAA 92
102 GCTGGGGTTCACTGGGTCACTTCCGCGAAGAGTGGCCATCAAGATCGTCAACCGTGAGAA 161
93 GATTGGAGAACATCAATTAAACAGGCCATAAAGTGGCAGTTTAAATCTTAAATAGACAGAA 152
162 GCTCAGCAGCTCGGTGCTGATG---AAGGTGGAGCGGGAGATCGCGATCCTGAAGCTCAT 218
153 GATTCCGAGTTTATGATGTTGTTGGAAAAATAAACAGAGAAATTCAAAATCTAAACTCTT 212
219 TGAGCACCCCACTGCTAAAGCTGCACGACGTTTATGAAAAAATAATTTGTACCT 278
213 TCCTCATCTCATATTATCAACTATATACAGGTGATCAGCACTCCACAGATTTTAT 272
279 GGTGCTAGAACACGTGTGAGTGTGAGTCTTTCGACTACTCGTGGTGAAGAGGGAGGCT 338
273 GGTATGGAAATATGTTGTGAGGTGAAATTAATTGACTATCATCTGTAAGCATGGACGGGT 332
339 GACGCTTAAGAGGCTCGGAAGTTCTTCGGCGAGATCATCTCGCGCTGCACTCTGCCA 398
333 TGAGAGATGGAAGCCAGGCGGTCTTTTACAGAGATCTCTGCTGTGATTTACTGTCA 392
399 CAGCCACTCCATATGCCACAGGGATCTGAAACCTGAAACCTCTGCTGGACGAGAGAA 458
393 TAGGCATATGTTGTTTCATCGAGACCTGAAACAGAGAAATGTCCTGTTGGATGCACACAT 452
459 CAACATCCGATCGCAGACTTTGGCATGGGCTCCCTGCAGGTTGGGCAAGCCCTGTTGA 518
453 GAATGCCAAGATAGCGGATTTCCGATTTATTAATATGATGTCAGATGGTGAATTTCTGAG 512
519 GACGAGCTGTGGTCCCGCCCACTACGCTCCCGAGGTGATCCGGGGGAGAGATGA 578
513 AACTAGTTCGGATCTCCAAATATATGACAGCACTGAAATCATCTCAGGCAAGATGTATGC 572
579 CGCCCGAAGCGGAGCTGTGGAGCTCGGCGTCACTCTGTTCCGCTTGTGTTGGGGG 638
573 AGGTCTGAAGTTGATATCTGGAGCTGTGGTGTATCTTGTATGCTCTCTTTGTGGCAC 632
639 TCTGCCCTTCGAGATGACAACTTTGCGACAGCTGTGGAGAGAGTGAAGCGGGCGGT 698
633 CCTCCCAATTTGATGATGAGCATGTACCTACGTTATTTAAGAGATCCGAGGGGTGCTT 692
699 CCATATCGCGACTTTATCCGCCCCGACTCCAGAGTCTGCTACGGGGCATGATCGAGGT 758
693 TTATATCCCAAGATATCTCAATCGTTCTGTGCGCACTCTCCTGATGATGTCGAGGT 752
759 GGAGCGCGCAGCGCGCTCACGCTAGAGCACATTCAGAAAAACATATGGTATATA 813

Db 753 TGACCCACTGAAACAGGCAACTATCAAGACATAAGAGAGCATGATGTTTAA 807
RESULT 13
US-08-557-006C-36
; Sequence 36, Application US/08557006C
; Patent No. 6258547
; GENERAL INFORMATION:
; APPLICANT: Berli, Rajindar K.
; APPLICANT: Carling, David
; APPLICANT: Forder, Robert A.
; TITLE OF INVENTION: NUCLEIC ACID ENCODING AMP-ACTIVATED PROTEIN KINASE
; FILE REFERENCE: NGAP/PHM37588/UST
; CURRENT APPLICATION NUMBER: US/08/557,006C
; CURRENT FILING DATE: 1996-03-06
; PRIOR APPLICATION NUMBER: PCT/GB94/01093
; PRIOR FILING DATE: 1994-05-20
; PRIOR APPLICATION NUMBER: GB 9310489.1
; PRIOR FILING DATE: 1993-05-21
; PRIOR APPLICATION NUMBER: GB 9318010.7
; PRIOR FILING DATE: 1993-08-31
; NUMBER OF SEQ ID NOS: 44
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 36
; LENGTH: 1783
; TYPE: DNA
; ORGANISM: Human AMP protein kinase
; US-08-557-006C-36
Query Match 8.8%; Score 176.6; DB 3; Length 1783;
Best Local Similarity 52.6%; Pred. No. 7.8e-31;
Matches 408; Conservative 0; Mismatches 364; Indels 3; Gaps 1;
QY 42 GTATGTTGGCCCTACCGCTGGAGAGACGCTGGGCAAGGGCGACAGAGTCTGTGAA 101
Db 34 GAAGATCGGACACTAGCTGCTGGGCGACACGCTGGGCGTGGCACCTTCGGCAAGTGAA 93
QY 102 GCTGGGGTTCACTGGGTCACTTCCGCGAAGAGTGGCCATCAAGATCGTCAACCGTGAGAA 161
Db 94 GATTGGAGAACATCAATTAAACAGGCCATAAAGTGGCAGTTTAAATCTTAAATAGACAGAA 153
QY 162 GCTCAGCAGCTCGGTGCTGATG---AAGGTGGAGCGGGAGATCGCGATCCTGAAGCTCAT 218
Db 154 GATTCCGAGTTTATGATGTTGTTGGAAAAATAAACAGAGAAATTCAAAATCTAAACTCTT 213
QY 219 TGAGCACCCCACTGCTAAAGCTGCACGACGTTTATGAAAAAATAATTTGTACCT 278
Db 214 TCCTCATCTCATATTATCAACTATATACAGGTGATCAGCACTCCACAGATTTTAT 273
QY 279 GGTGCTAGAACACGTGTGAGTGTGAGTCTTTCGACTACTCGTGGTGAAGAGGGAGGCT 338
Db 274 GGTATGGAAATATGTTGTGAGGTGAAATTAATTGACTATCATCTGTAAGCATGGACGGGT 333
QY 339 GACGCTTAAGAGGCTCGGAAGTTCTTCGGCGAGATCATCTCGCGCTGCACTCTGCCA 398
Db 334 TGAGAGATGGAAGCCAGGCGGTCTTTTACAGAGATCTCTGCTGTGATTTACTGTCA 393
QY 399 CAGCCACTCCATATGCCACAGGGATCTGAAACCTGAAACCTCTGCTGGACGAGAGAA 458
Db 394 TAGGCATATGTTGTTTCATCGAGACCTGAAACAGAGAAATGTCCTGTTGGATGCACACAT 453
QY 459 CAACATCCGATCGCAGACTTTGGCATGGGCTCCCTGCAGGTTGGCGACAGCTGTGGA 518
Db 454 GAATGCCAAGATAGCGGATTTCCGATTTATTAATATGATGTCAGATGGTGAATTTCTGAG 513
QY 519 GACGAGCTGTGGTCCCGCCCACTACGCTCCCGAGGTGATCCGGGGGAGAGATGA 578
Db 514 AACTAGTTCGGATCTCCAAATATATGACAGCACTGAAATCATCTCAGGCAAGATGTATGC 573
QY 579 CGCCCGAAGCGGAGCTGTGGAGCTCGGCGTCACTCTGTTCCGCTTGTGTTGGGGG 638
Db 574 AGGTCTGAAGTTGATATCTGGAGCTGTGGTGTATCTTGTATGCTCTCTTTGTGGCAC 633

639 TCTGCGCTTCGACGATGACAACTTCGCGACAGCTGCTGGAGAGAGTGAAGCGGGCGTGT 698
634 CTTCCCATTTGATGATGACATGTACTACGTATTATTAAAGAGATCCGAGGGGGTGTCTT 693
699 CCACATGCGCGACTTTATCCCGCCGAGCTGCGCAGAGTCTGCTACGGGGCATGTCGAGGT 758
694 TTATATCCAGAAATATCTCAATCGTTCTGTGCGCACTCTCTCTGATGATATGCTGCAGGT 753
759 GGACGCGGCACGCCCTCACGCTAGAGCACATTCAGAAACACATATGATATATA 813
754 TGACCCACTGAACAGCACTATCAAGACACATAGAGAGCATGAATGTTTAA 808

14
ence 1, Application US/08677298
nt No. 5863729
ERAL INFORMATION:
PPLICANT: Pionica-Worms, Helen
TITLE OF INVENTION: DNA SEQUENCES ENCODING HUMAN TcAK-1
TITLE OF INVENTION: KINASE
NUMBER OF SEQUENCES: 17
ORRESPONDENCE ADDRESS:
ADDRESSER: Greenlee, Winner and Sullivan, P.C.
STREET: 5370 Manhattan Circle, Suite 201
CITY: Boulder
STATE: CO
COUNTRY: USA
ZIP: 80303

OMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.30
URRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/677,298
FILING DATE: 09-JUL-1996
CLASSIFICATION: 536
TTORNEY/AGENT INFORMATION:
NAME: Caruthers, Jennie M.
REGISTRATION NUMBER: 34,464
REFERENCE/DOCKET NUMBER: 9-96
ELECOMMUNICATION INFORMATION:
TELEPHONE: (303) 499-8080
TELEFAX: (303) 499-8089
ORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 2698 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: not relevant
MOLECULE TYPE: cDNA to mRNA
HYPOTHETICAL: NO
ANTI-SENSE: NO
FEATURE:
NAME/KEY: CDS
LOCATION: 376..2565
-677-298-1

Match 8.4%; Score 168.2; DB 2; Length 2698;
Local Similarity 50.8%; Pred. No. 6.9e-29;
hes 401; Conservative 0; Mismatches 388; Indels 0; Gaps 0;
47 TTGGGCGCTTACCGGCTGGAGAAACGCTGGGCAAGGGGAGAGAGTCTGTTGAAGCTGG 106
533 TCGGAAACTACAGACTGTTGAAACCAATCGGCAAGGGGAATTTTGCAAAAGTAAATGG 592
107 GGGTTCTACTCGGTCACTGCGACAGAGTGGCCATCAAGATCGTCAACCGTGAGAGCTCA 166
593 CAAGACATATCTTTACAGGAGAGAGTTCATATAAAATAATGACAAACTCAGTTGA 652
167 GCGAGTGGTGTCTGATGAAGGTGGAGCGGGGAGATCGCGATCTCTGAAGCTCATTGAGCA 226

Db 653 ATCCAAACAAGTCTTACAAAAGCTCTTCAGAGAGTAAAGATAATGAAGATTTTAATCATC 712
Qy 227 CCCACGCTCTAAAGCTGACAGAGTCTTATGAAACAAAAATATTTTACCTGGTGTAG 286
Db 713 CCAATATAGTGAAGTTATTTCGAAGTCAATGAAACTGAAACAAACACTCTACTAATCATGG 772
Qy 287 AACAGTGTGAGTGTGAGCTCTTCGACTACCTGCTGAGAGAGAGGGAGGCTGACGGCTA 346
Db 773 AATATGCAAGTGGAGGTGAAGTATTTGACTATTTGCTGTCATGCGAGGATGAAGAAA 832
Qy 347 AGGAGGCTCGGAAGTCTTCCGCGAGATCATCTCTGCGTGGAGCTCTGCGCACAGCCACT 406
Db 833 AAGAGCAAGATCTAAATTTAGACAGATTTGCTGCACTCAATATCTGCTCAGAAAC 892
Qy 407 CCATATGCCACAGGGATCTGAAACCTGAAACCTCTGCTGGAGAGAGAAACAAATCC 466
Db 893 GGATCGTACATCGAGACTCAAGGCTGAAATCTTATTGTTAGATGCCGATATGAACATTA 952
Qy 467 GCATCGCAGACTTTGGCATGGCGTCCCTGAGGTTGCGGACAGCTGTTGGAGACAGCT 526
Db 953 AATATGACAGATTTTCGGTTTATAGCAATGAAATTTACTGTTGGCGGTAAACTCGACACGTTT 1012
Qy 527 GTGGTCCCGCCACTACGCTGCCCCGAGGTGATCCGGGGGAGAGTATGACGCGCGGA 586
Db 1013 GTGCGAGTCTCCATACGACGACCTGAGCTCTTCAGGGGCAAGAAATATGACGGGCGAG 1072
Qy 587 AGCGGAGCTGTGAGCTGGCGCTCATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 646
Db 1073 AAGTGGATGTGAGGCTGGGGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1132
Qy 647 TCGACGATGACAACTTCGACAGCTGCTGGAGAGGTGAAGCGGGGCGTGTTCACATGC 706
Db 1133 TTGATGGGCAAACTTAAAGAACTGAGAGAGAGAGTATTAAAGAGGGGAAATACAGAAATC 1192
Qy 707 CGCACTTTATCCCGCCGACTGCGAGAGTCTGCTACGCGGCGATGATCGAGTGGACGCG 766
Db 1193 CTTCTACATGCTTACAGACTGTGAAACCTTCTCAAACGTTTCTGCTGCTAAATCCAA 1252
Qy 767 CACGCGCTCACGCTAGAGACATTCAGAAACACATATGTTATATAGGGGGCAGAAATG 826
Db 1253 TTAACGCGGCACTTAGAGCAAAATCATGAAGGACAGGTGGATCAATGAGGGCATGAAG 1312
Qy 827 AGCCCGAC 835
Db 1313 AAGATGAAC 1321

RESULT 15
US-09-523-849-4
Sequence 4, Application US/09523849
Patent No. 6458561
GENERAL INFORMATION:
APPLICANT: Bandman, Olga
APPLICANT: Molteni, Angela
APPLICANT: Magnaghi, Paola
APPLICANT: Bobotti, Roberta
APPLICANT: Scacheri, Emanuela
APPLICANT: Isacchi, Antonella
APPLICANT: Hodgson, Dave
TITLE OF INVENTION: HUMAN NIM1 KINASE
FILE REFERENCE: PC-0009 US
CURRENT APPLICATION NUMBER: US/09/523,849
CURRENT FILING DATE: 2000-03-13
NUMBER OF SEQ ID NOS: 39
SOFTWARE: PERL Program
SEQ ID NO 4
LENGTH: 1507
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: misc feature
OTHER INFORMATION: Incyte Clone No. 6458561 670279CB1

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

ucleic - nucleic search, using sw model

in: November 27, 2003, 01:57:44 ; Search time 456.512 Seconds
(without alignment)
10603.388 Million cell updates/sec

US-10-054-579-3

ct score: 1827

nce: 1. atgaaggtggagcggagat.....gaattatccgaaagtaa 1827

ng table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

hed: 2552756 seqs, 1349719017 residues

. number of hits satisfying chosen parameters: 5105512

um DB seq length: 0

um DB seq length: 2000000000

processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

ase : N Geneseq 19Jun03.*

- 1: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1980.DAT.*
- 2: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1981.DAT.*
- 3: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1982.DAT.*
- 4: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1983.DAT.*
- 5: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1984.DAT.*
- 6: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1985.DAT.*
- 7: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1986.DAT.*
- 8: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1987.DAT.*
- 9: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1988.DAT.*
- 10: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1989.DAT.*
- 11: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1990.DAT.*
- 12: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1991.DAT.*
- 13: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1992.DAT.*
- 14: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1993.DAT.*
- 15: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1994.DAT.*
- 16: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1995.DAT.*
- 17: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1996.DAT.*
- 18: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1997.DAT.*
- 19: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1998.DAT.*
- 20: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT.*
- 21: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT.*
- 22: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT.*
- 23: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT.*
- 24: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2002.DAT.*
- 25: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2003.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

t	Score	Query Match	Length	DB ID	Description
1	1827	100.0	1827	24	ABQ77626 Human kinase NHP #
2	1827	100.0	2007	24	ABQ77625 Human kinase NHP #
3	1827	100.0	2647	24	AAD34315 Human PKIN-18 cDNA
4	1825.4	99.9	2908	25	ABX13583 Human cDNA encoding
5	1813	99.2	3364	25	ABX13584 Human cDNA encoding
6	1759.6	96.3	2025	24	ABA02995 Human protein kina
7	1759.6	96.3	2217	24	ABA02994 Human protein kina
8	904.2	49.5	2719	25	AAL51889 Human cell cycle-r

9	851.4	46.6	2112	25	ABV74559	Human kinase #3 co
10	851.4	46.6	2337	25	ABV74557	Human kinase #1 co
11	851.4	46.6	2385	22	AAS06717	Polynucleotide seq
12	851.4	46.6	2897	24	AAD26464	Human kinase PKIN-
13	777.2	42.5	2289	25	ABV74558	Human kinase #2 co
14	598.8	32.2	1797	25	ABX70994	Novel human cDNA s
15	446.2	24.4	2720	23	ABL10489	Drosophila melanog
16	393.8	21.6	906	22	ABA08296	Human HrPopK-i hom
17	291.6	16.0	512	24	ABK70216	Human lung cancer
18	249.8	13.7	614	22	ABA08924	Human Ser/thr kina
19	232.4	12.7	1594	22	AAF46555	Novel protein kina
20	229.8	12.6	3170	24	ABA05737	Murine neuronal se
21	229.8	12.6	3250	24	ABA05738	Murine neuronal se
22	225	12.3	1549	24	ABG64386	Human serine/threo
23	225	12.3	2085	25	AAD51409	Human microtubule
24	225	12.3	2278	25	AAD51410	Human microtubule
25	225	12.3	2462	24	AQ93445	Human cDNA SEQ ID
26	225	12.3	2954	22	AAK94923	Human full-length
27	225	12.3	3226	25	AAD51412	Human microtubule
28	225	12.3	3269	22	AAD18831	Human kinase (PKIN
29	225	12.3	3270	25	ABQ77242	Human MARK cDNA.
30	225	12.3	3312	24	ABA05739	Human neuronal ser
31	225	12.3	3392	24	ABA05740	Human neuronal ser
32	225	12.3	3609	25	AAD51408	Human microtubule
33	225	12.3	4917	25	AAD51411	Human microtubule
34	223	12.2	2283	25	ABZ77146	Human protein kina
35	222.8	12.2	2352	22	AAI66823	Human protein kina
36	222.8	12.2	2533	21	AAK82951	Human keratinocyte
37	222.8	12.2	2632	21	AAK82952	Human keratinocyte
38	222.8	12.2	2968	22	AAI66822	Human protein kina
39	222.8	12.2	4699	22	AAH76213	Human kinase PKIN-
40	218.6	12.0	2663	23	ABU05573	Drosophila melanog
41	206.6	11.3	859	22	AAK91635	Human cDNA 5'-end
42	206.6	11.3	859	22	AAK93561	Human cDNA clone r
43	205.6	11.3	1349	22	AAI99637	Human expressed po
44	205.6	11.3	1349	23	AAK43708	DNA encoding novel
45	203.8	11.2	2361	22	AAS06718	Polynucleotide seq

ALIGNMENTS

RESULT 1

ABQ77626

ID ABQ77626 standard; cDNA; 1827 BP.

XX ABQ77626;

AC ABQ77626;

XX 21-OCT-2002 (first entry)

DT Human kinase NHP #2-encoding cDNA, SEQ ID NO:3.

DE Human kinase NHP; chromosome 11; signal transduction; cytoplasmic;

XX Human; kinase; NHP; chromosome 11; signal transduction; cytoplasmic;

KW nuclear; brain; pituitary; hypothalamus; adipose; cerebellum;

KW adrenal gland; foetal lung; foetal brain; cancer; drug screening;

KW mental disorder; diagnostic reagent; clinical trial monitoring; cosmetic;

KW nutraceutical; mutation detection; gene expression analysis;

KW chromosome mapping; transgenic animal; cycostatic; gene therapy;

XX gene; ss.

XX Homo sapiens.

OS OS

XX Key

XX Location/Qualifiers

FT CDS

FT 1..1827

FT /*tag= a

FT /product= "Human kinase NHP #2"

FT WO200259287-A2.

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

23-JAN-2001; 2001US-263378P.

(LEXI-) LEXICON GENETICS INC.

Turner CA, Mathur B;

4PI; 2002-599780/64.

P-PSDB; ABB09558.

Novel polynucleotide encoding human proteins sharing sequence similarity with animal kinases, useful for drug screening, diagnosis, in gene therapy of disorders and diseases e.g. cancer

Claim 3; Page 38-39; 40pp; English.

The invention relates to novel human kinases (ABB09557, ABB09558), referred to as NHPs, and to nucleic acids encoding them (ABB07625, ABB07626). The NHPs have structural similarity with animal kinases including serine-threonine kinases, carbon catabolite repressing kinases, calcium/calmodulin-dependent protein kinases, and cAMP-dependent protein kinases. The NHPs do not contain consensus signal sequences, indicating that they may be cytoplasmic or nuclear proteins, and are thought to be involved in signal transduction. Polynucleotides encoding NHPs were obtained using human brain, pituitary, hypothalamus, adipose, cerebellum, adrenal gland, foetal lung, and foetal brain cDNAs and primers derived from human genomic DNA. The gene encoding the NHPs is located on chromosome 11. NHP nucleotides and proteins are useful for treating mental, biological or medical disorders including cancer, and for screening compounds useful in the treatment of such conditions. They can also be used as diagnostic reagents, in clinical trial monitoring and in cosmetic and nutritional applications. NHP nucleotides can additionally be used in the detection of disease-associated mutations, in the analysis of gene expression, for mapping chromosome 11, for the recombinant expression of NHPs, to generate transgenic animals, in gene therapy, and as part of ribozyme and/or triple helix sequences useful in the modulation of NHP gene expression. The present sequence represents cDNA encoding a 608 amino acid NHP (#2).

Sequence 1827 BP; 406 A; 581 C; 535 G; 305 T; 0 other;

CY Match 100.0%; Score 1827; DB 24; Length 1827;
Local Similarity 100.0%; Pred. No. 0;
Matches 1827; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 ATGAAGTGGAGCGGAGATCGGATCCTGAAGCTCATTTAGACACCCCGCCTCTAAAG 60
1 ATGAAGTGGAGCGGAGATCGGATCCTGAAGCTCATTTAGACACCCCGCCTCTAAAG 60
61 CTGCACGACGTTTATGAACAACAAATATTTGTAACCTGTGTAGAACACGTTTCAGGT 120
61 CTGCACGACGTTTATGAACAACAAATATTTGTAACCTGTGTAGAACACGTTTCAGGT 120
121 GGTGAGCTCTTCCGACTACCTGTTGAAGAGGAGGCTGACGCTTAAGGAGGCTCGGAAG 180
121 GGTGAGCTCTTCCGACTACCTGTTGAAGAGGAGGCTGACGCTTAAGGAGGCTCGGAAG 180
181 TTCTTCGGCAGATCATCTCTGGCTGAGCTTTCGACAGCCATTCATATGCCACAGG 240
181 TTCTTCGGCAGATCATCTCTGGCTGAGCTTTCGACAGCCATTCATATGCCACAGG 240
241 GATCTGAACCTGAAACCTCTGCTGGACGAGAGAACACATCCGATCGCAGACTTT 300
241 GATCTGAACCTGAAACCTCTGCTGGACGAGAGAACACATCCGATCGCAGACTTT 300
301 GGATGAGCTCCCTGACAGGTTGGGACAGCTGTTGGAGACCACTGTGGGTCCCCCAC 360
301 GGATGAGCTCCCTGACAGGTTGGGACAGCTGTTGGAGACCACTGTGGGTCCCCCAC 360
361 TAGCCCTGCCCGAGGTGATCCGGGGGAGAGATGACCGCCGAGAGCGGAGCTGTGG 420
361 TAGCCCTGCCCGAGGTGATCCGGGGGAGAGATGACCGCCGAGAGCGGAGCTGTGG 420
421 AGCTGGCGGCTCATCTCTGTTCCGCTTCTGTTGGGGGCTCTGCCCTTCGACGATGACAC 480

421 AGCTGGCGGCTCATCTCTGTTCCGCTTGTGTTGGGGCTCTGCCCTTCGACGATGACAAAC 480
481 TTGGACAGCTGCTGGAGAGGTGAAGCGGGCGGTGTTCCACATGCGGCACTTTATCCGG 540
481 TTGGACAGCTGCTGGAGAGGTGAAGCGGGCGGTGTTCCACATGCGGCACTTTATCCGG 540
541 CCCGACTCCAGAGTCTCTACGGGGCATGATCGAGGTGGAGCGCGCACGCGCCTCAGG 600
541 CCCGACTCCAGAGTCTCTACGGGGCATGATCGAGGTGGAGCGCGCACGCGCCTCAGG 600
601 CTAGAGCAGCTTCAGAAACACATATGTTATAGGGGCAAGATGAGCCCGAACAGAG 660
601 CTAGAGCAGCTTCAGAAACACATATGTTATAGGGGCAAGATGAGCCCGAACAGAG 660
661 CAGCCCATCTCTCGCAAGGTGCAGATCCGCTCGCTGCCCGCAGCTGGAGGACATGACCCC 720
661 CAGCCCATCTCTCGCAAGGTGCAGATCCGCTCGCTGCCCGCAGCTGGAGGACATGACCCC 720
721 GACGTGCTGGAGCAGCATGCACTCACTGGGCTGCTTCCGAGACCGCAACAAAGTGTGCGAG 780
721 GACGTGCTGGAGCAGCATGCACTCACTGGGCTGCTTCCGAGACCGCAACAAAGTGTGCGAG 780
781 GACGTGCTGGAGGAGAGAACACAGAGAGATGATTTACTTCTCTCTCTGAGACCGG 840
781 GACGTGCTGGAGGAGAGAACACAGAGAGATGATTTACTTCTCTCTCTGAGACCGG 840
841 AAAGAAAGTACCCGAGCAGGAGATGAGGACCTGCCCCCGGAGCAGATGACCCCT 900
841 AAAGAAAGTACCCGAGCAGGAGATGAGGACCTGCCCCCGGAGCAGATGACCCCT 900
901 CCCCGAAGCGTGTGGACTCCCGGATGTGAACCGGACGCGGCGGCGGCGCAAGACGC 960
901 CCCCGAAGCGTGTGGACTCCCGGATGTGAACCGGACGCGGCGGCGGCGCAAGACGC 960
961 AAATCCATGGAGGTGCTCAGCGTGAAGCGGCGCTCCCGGTGCTGCGGGCGGGCC 1020
961 AAATCCATGGAGGTGCTCAGCGTGAAGCGGCGCTCCCGGTGCTGCGGGCGGGCC 1020
1021 ATTGAGATGGCCAGCGGCGGAGGCTCGGTCCATCAGCGGTGCTCTCAGGCGCTT 1080
1021 ATTGAGATGGCCAGCGGCGGAGGCTCGGTCCATCAGCGGTGCTCTCAGGCGCTT 1080
1081 TCCACAGCCCACTCAGCAGCGCCCGGGTGAGCCCTCACCCCTCACAAAGGGCAGTCCC 1140
1081 TCCACAGCCCACTCAGCAGCGCCCGGGTGAGCCCTCACCCCTCACAAAGGGCAGTCCC 1140
1141 CTCGCCACCCCGGAGGACCTGTCACACGCGCAAGGAGAGCGCGCTGGCAGCGCC 1200
1141 CTCGCCACCCCGGAGGAGACCTGTCACACGCGCAAGGAGAGCGCGCTGGCAGCGCC 1200
1201 AACCCAGCGCCCGCTCAGCGCCAGCGCTCGGAGGGGTGCCCTGGAGGCGCGGTCAAC 1260
1201 AACCCAGCGCCCGCTCAGCGCCAGCGCTCGGAGGGGTGCCCTGGAGGCGCGGTCAAC 1260
1261 TCCATCAAGAACAGCTTTCTGGGCTCACCCCGCTTCCACCGCCGGAATCTGCAAGTTCG 1320
1261 TCCATCAAGAACAGCTTTCTGGGCTCACCCCGCTTCCACCGCCGGAATCTGCAAGTTCG 1320
1321 AGCCCGGAGGAGATGTCCTCAACCTGACACGAGTCTGCCCGAGCTGGCGAGAGTCC 1380
1321 AGCCCGGAGGAGATGTCCTCAACCTGACACGAGTCTGCCCGAGCTGGCGAGAGTCC 1380
1381 TGGTTGGGAACCTTCATCAGCTCGAGAGGAGGAGAGATCTTCTGTTGTTCTCAAGAGAC 1440
1381 TGGTTGGGAACCTTCATCAGCTCGAGAGGAGGAGAGATCTTCTGTTGTTCTCAAGAGAC 1440
1441 AAACCTCTGAGCTCCATCAAGGCTGACATCGTGCACGCTTCTCTGATTCCTCAGTCTC 1500
1441 AAACCTCTGAGCTCCATCAAGGCTGACATCGTGCACGCTTCTCTGATTCCTCAGTCTC 1500
1501 AGCCACAGGCTCATCTCCCAACAGCTTCCGGGCGGAGTACAGGCCACCGGGGGGCA 1560

721 GACGTGCTGGACAGATGCACTCACTGGGCTGCTTCGAGAGCCGCAACAAGCTGCTGCAG 780
901 GACGTGCTGGACAGATGCACTCACTGGGCTGCTTCGAGAGCCGCAACAAGCTGCTGCAG 960
781 GACGTGCTGCTCCGAGGAGGAGAACACGAGAGAGATGATTACTTCTCTCTCTGGACCGG 840
961 GACGTGCTGCTCCGAGGAGGAGAACACGAGAGAGATGATTACTTCTCTCTCTGGACCGG 1020
841 AAGAAGAGGTACCCGAGCAGGAGGATGAGACCTGCCCCCGGAGACGAGATAGACCTT 900
1021 AAGAAGAGGTACCCGAGCAGGAGGATGAGACCTGCCCCCGGAGACGAGATAGACCTT 1080
901 CCCCGAGAGGTGAGGACTCCCGATCTGACCGGACGAGCAGCGCGGCGCCAGAACGC 960
1081 CCCCGAGAGGTGAGGACTCCCGATCTGACCGGACGAGCAGCGCGGCGCCAGAACGC 1140
961 AATTCATGAGGTGCTCAGCGTACGAGCGCGGCTCCCGGCTGCTGCGCGCGCGGCC 1020
1141 AATTCATGAGGTGCTCAGCGTACGAGCGCGGCTCCCGGCTGCTGCGCGCGGCC 1200
1021 ATTGAGATGGCCAGCAGCGGCTCAGGTCTCGGTCCATCAGCGGTGCTCTCTCAGGCTT 1080
1201 ATTGAGATGGCCAGCAGCGGCTCAGGTCTCGGTCCATCAGCGGTGCTCTCTCAGGCTT 1260
1081 TCCACCGCCCACTCAGCAGCGGCTGAGCGGCTCAGCGGCTCAGCGGCTCAGCGGCTT 1140
1261 TCCACCGCCCACTCAGCAGCGGCTGAGCGGCTCAGCGGCTCAGCGGCTCAGCGGCTT 1320
1141 CTCCCGCAGCGGCTGAGCGGCTCAGCGGCTCAGCGGCTCAGCGGCTCAGCGGCTT 1200
1321 CTCCCGCAGCGGCTGAGCGGCTCAGCGGCTCAGCGGCTCAGCGGCTCAGCGGCTT 1380
1201 AACCCAGCGGCTGAGCGGCTCAGCGGCTCAGCGGCTCAGCGGCTCAGCGGCTT 1260
1381 AACCCAGCGGCTGAGCGGCTCAGCGGCTCAGCGGCTCAGCGGCTCAGCGGCTT 1440
1261 TCCATCAGAGCACTTTCGCGCTCAGCGGCTCAGCGGCTCAGCGGCTCAGCGGCTT 1320
1441 TCCATCAGAGCACTTTCGCGCTCAGCGGCTCAGCGGCTCAGCGGCTCAGCGGCTT 1500
1321 ACGCCGAGGAGATGTCACACCTGACACGAGTCTGCTCCCGAGAGCTGGGAGAGAGTCC 1380
1501 ACGCCGAGGAGATGTCACACCTGACACGAGTCTGCTCCCGAGAGCTGGGAGAGTCC 1560
1381 TGGTTTGGAACTTTCATCAGCTGAGAGAGAGAGAGAGATTTCTGTTGATCAAAAGAC 1440
1561 TGGTTTGGAACTTTCATCAGCTGAGAGAGAGAGAGAGATTTCTGTTGATCAAAAGAC 1620
1441 AACCTCTGAGCTCCATCAGGCTGACATCGTGACGCTTCTGTCGATTTCCAGTCTC 1500
1621 AACCTCTGAGCTCCATCAGGCTGACATCGTGACGCTTCTGTCGATTTCCAGTCTC 1680
1501 AGCCACAGCGTCTATCTCCAAACGAGCTTCCGCGCCGAGTACAAAGGCGCGGCGCCA 1560
1681 AGCCACAGCGTCTATCTCCAAACGAGCTTCCGCGCCGAGTACAAAGGCGCGGCGCCA 1740
1561 GCGGTGTTCCAGAGCGGCTCAAGTTCAGGTTGATATCAGCTTCTGTCGATTTCCAGTCTC 1620
1741 GCGGTGTTCCAGAGCGGCTCAAGTTCAGGTTGATATCAGCTTCTGTCGATTTCCAGTCTC 1800
1621 GCGCAGAGGAGAGCGGCTATCTCTGTCACCTTCCAGCTGCTCTCAGGCGCCAGCGCT 1680
1801 GCGCAGAGGAGAGCGGCTATCTCTGTCACCTTCCAGCTGCTCTCAGGCGCCAGCGCT 1860
1681 CGCTTCCAGAGGCGGCTGAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1740
1861 CGCTTCCAGAGGCGGCTGAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1920
1741 GCGGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1800
1921 GCGGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1980
1801 AAATGTGGAATTATCCCGAAAGTTAA 1827

Db 1981 AAATGTGGAATTATCCCGAAAGTTAA 2007
RESULT 3
AAD34315
ID AAD34315 standard; cDNA; 2647 BP.
XX AAD34315;
XX AC
XX 16-JUL-2002 (first entry)
XX Human PKIN-18 cDNA.
XX Human; kinase; enzyme; PKIN-18 protein; immune system disorder; anaemia;
KW acquired immune deficiency syndrome; thymic hypoplasia; Crohn's disease;
KW asthma; neurological disorder; epilepsy; Charcot-Marie-Tooth disease;
KW AIDS; seizures; cell proliferative disorder; cancer; adenocarcinoma;
KW leukaemia; lymphoma; melanoma; sarcoma; developmental disorder;
XX Down's syndrome; gene therapy; protein therapy; cytostatic; gene; ss.
XX Homo sapiens.
OS
PH Key Location/Qualifiers
FT 1..1995
FT /*tag= a
FT /product= "Human PKIN-18 protein"
XX WO200218557-A2.
PD 07-MAR-2002.
XX 31-AUG-2001; 2001WO-US27219.
XX 31-AUG-2000; 2000US-329873P.
PR 08-SEP-2000; 2000US-231357P.
PR 14-SEP-2000; 2000US-232854P.
PR 22-SEP-2000; 2000US-234902P.
PR 29-SEP-2000; 2000US-236499P.
PR 06-OCT-2000; 2000US-238389P.
PR 13-OCT-2000; 2000US-240542P.
XX (INCY-) INCYTE GENOMICS INC.
XX Bandman O, Nguyen DB, Walia NK, Hafalia AJA, Yao MG, Gandhi AR;
PI Gururajan R, Ding L, Patterson C, Yue H, Baughn MR, Tribouley CM;
PI Thornton M, Elliott VS, Lu Y, Ison CH, Au-Young J, Tang YT;
PI Azimzai Y, Burrill JD, Marcus GA, Zingler KA, Lu DAM, Lal PG;
PI Ramkumar J, Warren BA, Kearney L, Policky JL, Thangavelu K;
PI Burford N;
XX WPI; 2002-329769/36.
DR P-PSDB; AAE21723.
XX New human kinases, useful for diagnosing, treating or preventing immune
PT system disorders (e.g. Crohn's disease), neurological disorders (e.g.
PT epilepsy), or cell proliferative disorders (e.g. cancers such as
PT leukemia or lymphoma)
XX Claim 97; Page 212-213; 218pp; English.
XX The present invention relates to human kinases (PKIN) and polynucleotides
CC encoding such proteins. PKIN sequences of the invention are useful for
CC diagnosing, treating or preventing disorders associated with aberrant
CC expression of PKIN, particularly immune system disorders (e.g. acquired
CC immune deficiency syndrome (AIDS), thymic hypoplasia, Crohn's disease,
CC anaemia, asthma), neurological disorders (e.g. epilepsy, Charcot-Marie-
CC tooth disease or seizures), cell proliferative disorders (e.g. cancers
CC such as adenocarcinoma, leukaemia, lymphoma, melanoma, myeloma, sarcoma),
CC and developmental disorders (e.g. Down's syndrome). They are also used
CC in gene therapy and protein therapy. The present sequence is a cDNA
CC encoding human PKIN-18 protein.

SQ	Sequence 2647 BP; 525 A; 885 C; 789 G; 448 T; 0 other;	
	Query Match 100.0%; Score 1827; DB 24; Length 2647;	
	Best Local Similarity 100.0%; Pred. No. 0;	
	Matches 1827; Conservative 0; Mismatches 0; Indels 0; Gaps 0;	
QY	1 ATGAAGGTGGAGCGGAGATCGCATCTCTGAAGCTCATTTGAGCACCCGCCCAAGTCTCTAAG 60	
DB	169 ATGAAGGTGGAGCGGAGATCGCATCTCTGAAGCTCATTTGAGCACCCGCCCAAGTCTCTAAG 228	
QY	61 CTGACGAGCTTTATGAAGAAACAAATATTTGTACCTGTGTGCTAGAACACGTTGTCAGGT 120	
DB	229 CTGACGAGCTTTATGAAGAAACAAATATTTGTACCTGTGTGCTAGAACACGTTGTCAGGT 288	
QY	121 GGTGAGCTCTTCCAGTACCTGTGTGAAGAGGGAGGCTGAGCCCTTAAGAGGCTCGGAAG 180	
DB	289 GGTGAGCTCTTCCAGTACCTGTGTGAAGAGGGAGGCTGAGCCCTTAAGAGGCTCGGAAG 348	
QY	181 TTCTTCCGCGAGATCATCTCTGCGTGGACTTTCTGCCACAGCACTCCATATGCCACAGG 240	
DB	349 TTCTTCCGCGAGATCATCTCTGCGTGGACTTTCTGCCACAGCACTCCATATGCCACAGG 408	
QY	241 GATCTGAACCTGAAACCTCTCTGTGACGAGAGAAACACATCCGATCGCAGACTTT 300	
DB	409 GATCTGAACCTGAAACCTCTCTGTGACGAGAGAAACACATCCGATCGCAGACTTT 468	
QY	301 GGCATGGCGTCTCCAGGTTGGCGACAGCCTGTTGGAGACAGCTGTGGTCCGCCAC 360	
DB	469 GGCATGGCGTCTCCAGGTTGGCGACAGCCTGTTGGAGACAGCTGTGGTCCGCCAC 528	
QY	361 TACGCTGCCCCAGGTTGATCCGGGGGAGAAATGATGACGGCCGGAAGCGGACGTGTGG 420	
DB	529 TACGCTGCCCCAGGTTGATCCGGGGGAGAAATGATGACGGCCGGAAGCGGACGTGTGG 588	
QY	421 AGCTGGGGCTCATCTGTTGCCCTTGTGTGGGGGCTCTGCCCTTCGACGATGACAAAC 480	
DB	589 AGCTGGGGCTCATCTGTTGCCCTTGTGTGGGGGCTCTGCCCTTCGACGATGACAAAC 648	
QY	481 TTGCGACAGCTGCTGAGAGAGTGAAGCGGGGCGTGTTCACATGCGCGCACTTTATCCCG 540	
DB	649 TTGCGACAGCTGCTGAGAGAGTGAAGCGGGGCGTGTTCACATGCGCGCACTTTATCCCG 708	
QY	541 CCGGACTGCCAGAGTCTGCTACGGGGCATATGAGAGTGGAGCGCGCACGCCGCTCAGC 600	
DB	709 CCGGACTGCCAGAGTCTGCTACGGGGCATATGAGAGTGGAGCGCGCACGCCGCTCAGC 768	
QY	601 CTAGAGCACATTCAGAAACACATATGTTATAGGGGGCAGAAATGAGCCCGAACACAG 660	
DB	769 CTAGAGCACATTCAGAAACACATATGTTATAGGGGGCAGAAATGAGCCCGAACACAG 828	
QY	661 CAGCCCATTTCTCGCAAGGTGCAGATCCGCTCGCTGCGCCAGCCTGGAGGACATCGACCCC 720	
DB	829 CAGCCCATTTCTCGCAAGGTGCAGATCCGCTCGCTGCGCCAGCCTGGAGGACATCGACCCC 888	
QY	721 GACGTGCTGGACAGATGACATCTATGGGCTGCTTCCGAGACCGCAACAAAGCTGTGCAG 780	
DB	889 GACGTGCTGGACAGATGACATCTATGGGCTGCTTCCGAGACCGCAACAAAGCTGTGCAG 948	
QY	781 GACCTGCTGTCGAGGAGGAGAACACAGAGAGATGATTTACTTCTCTCTCTGACCGG 840	
DB	949 GACCTGCTGTCGAGGAGGAGAACACAGAGAGATGATTTACTTCTCTCTCTGACCGG 1008	
QY	841 AAGAAAGGTATCCGAGCGAGGATGAGGACCTGCGCCCGGAAACGAGATAGACCTT 900	
DB	1009 AAGAAAGGTATCCGAGCGAGGATGAGGACCTGCGCCCGGAAACGAGATAGACCTT 1068	
QY	901 CCCCAGGAGCTGTGGACTTCCCGCATGTCTAAACCGGCACGGCAAGCGCGCGCAGAAACGC 960	
DB	1069 CCCCAGGAGCTGTGGACTTCCCGCATGTCTAAACCGGCACGGCAAGCGCGCGCAGAAACGC 1128	
QY	961 AATCCATGAGGCTCTCAGCTGACGACGCGGCTCCCGGCTCCCTGCGGGGGGCC 1020	
DB	1129 AATCCATGAGGCTCTCAGCTGACGACGCGGCTCCCGGCTCCCTGCGGGGGGCC 1188	

QY	1021 ATTGAGATGCCCCAGCACGCGCCAGAGGTCTCGTCCATCAGCGTGCCTCTCAGGCGCTT 1080	
DB	1189 ATTGAGATGCCCCAGCACGCGCCAGAGGTCTCGTCCATCAGCGTGCCTCTCAGGCGCTT 1248	
QY	1081 TCACACGACCCACTCAGCAGCCCCCGGGTGACCCCTCACCCCTCACCAAGGGGAGTCCC 1140	
DB	1249 TCACACGACCCACTCAGCAGCCCCCGGGTGACCCCTCACCCCTCACCAAGGGGAGTCCC 1308	
QY	1141 CTCCCCACCCCAAGGGGACACCTGTCCACGCGCAAGAGAGCCCGGCTGGGACGCGCC 1200	
DB	1309 CTCCCCACCCCAAGGGGACACCTGTCCACGCGCAAGAGAGCCCGGCTGGGACGCGCC 1368	
QY	1201 AACCCCAAGCCCCCGCTCCAGCCCCAGCGGTGGAGGGGTGCCCTGGAGGGCGCGCTCAAC 1260	
DB	1369 AACCCCAAGCCCCCGCTCCAGCCCCAGCGGTGGAGGGGTGCCCTGGAGGGCGCGCTCAAC 1428	
QY	1261 TCCATCAAGAACAGCTTTTGGGCTCACCCCGCTTCCACCGCCGGAACCTCAAGTTCCG 1320	
DB	1429 TCCATCAAGAACAGCTTTTGGGCTCACCCCGCTTCCACCGCCGGAACCTCAAGTTCCG 1488	
QY	1321 ACGCGGAGGAGATGTCACACCTGACACCGAGTCTGCCCGAGAGCTGGGCGAAGTCCC 1380	
DB	1489 ACGCGGAGGAGATGTCACACCTGACACCGAGTCTGCCCGAGAGCTGGGCGAAGTCCC 1548	
QY	1381 TGGTTTGGGAACTTCATCAGCTCGGAGAGGAGGAGAGATCTTCGTGGTTCATCAAGAGAC 1440	
DB	1549 TGGTTTGGGAACTTCATCAGCTCGGAGAGGAGGAGAGATCTTCGTGGTTCATCAAGAGAC 1608	
QY	1441 AAACCTCTGAGCTCCATCAAGGCTGACATCGTGCAACGCTTCTGTGATTTCCAGTCTC 1500	
DB	1609 AAACCTCTGAGCTCCATCAAGGCTGACATCGTGCAACGCTTCTGTGATTTCCAGTCTC 1668	
QY	1501 AGCCACAGCTCATCTCCCAACGAGCTTCGGGCGAGGTACAGGCCACGGGGGGGCA 1560	
DB	1669 AGCCACAGCTCATCTCCCAACGAGCTTCGGGCGAGGTACAGGCCACGGGGGGGCA 1728	
QY	1561 GCCGTGTTCCAGAGCGGTCAAGTTCAGGTTGATATCACTACACGGAGGCTGGGGAG 1620	
DB	1729 GCCGTGTTCCAGAGCGGTCAAGTTCAGGTTGATATCACTACACGGAGGCTGGGGAG 1788	
QY	1621 GCGCAGAGGAGAACCGGATCTACTCCGTCACTTCACTTCCCTGTCTCAGGCCCCAGCGGT 1680	
DB	1789 GCGCAGAGGAGAACCGGATCTACTCCGTCACTTCACTTCCCTGTCTCAGGCCCCAGCGGT 1848	
QY	1681 CGCTTCAAGAGGGTGGAGACCATCCAGGCCAGCTGTGAGCACACACGACCCGCT 1740	
DB	1849 CGCTTCAAGAGGGTGGAGACCATCCAGGCCAGCTGTGAGCACACACGACCCGCT 1908	
QY	1741 GCGGCCAGCATTTGTGAGACACCACTAATCTGTATGAAATGATGACGGGCGGCTTTCC 1800	
DB	1909 GCGGCCAGCATTTGTGAGACACCACTAATCTGTATGAAATGATGACGGGCGGCTTTCC 1968	
QY	1801 AATGTGGAATTTATCCGAAAAGTTAA 1827	
DB	1969 AATGTGGAATTTATCCGAAAAGTTAA 1995	

RESULT 4
ABX13583
ID ABX13583 standard; cDNA; 2908 BP.
XX
AC ABX13583;
XX
XX
DT 10-FEB-2003 (first entry)
XX
DE Human cDNA encoding serine protein kinase KSE336-1.
XX
KW Human; ss; gene; chromosome 11p15.5-pter; astrocytoma; meningioma;
KW pancreatic adenocarcinoma; insulin-dependent diabetes mellitus 2;
KW helioid peripapillary choriorretinal degeneration; brain; pancreas;
KW Beckwith-Wiedemann syndrome; congenital hyperinsulinism; KSE336.
XX

OS	Homo sapiens.	Key	Location/Qualifiers	DB
XX		PH	106..2112	585
FT		FT	/*tag= a	360
FT		FT	/product= "KSE336-1"	645
XX				420
XX	US6455292-B1.			705
XX	24-SEP-2002.			480
XX				765
XX	16-AUG-2001; 2001US-0930181.			540
XX	16-AUG-2001; 2001US-0930181.			825
XX	(ORIG-) ORIGENE TECHNOLOGIES INC.			600
XX	Shu Y, Fan W, Kovacs KF, Zidanic M, Jay G;			885
XX	WPI; 2003-066233/06.			660
XX	P-PSDB; ABG72382.			945
XX	New isolated polynucleotide coding without interruption for a human			720
XX	KSE336 polypeptide useful for preventing or treating			1005
XX	diseases/conditions relating to brain and pancreas, e.g. meningioma,			780
XX	insulin-dependent diabetes mellitus 2			1065
XX	Claim 3; Column 33-40; 34pp; English.			840
XX	The invention relates to an isolated polynucleotide (its complement			1125
XX	or a sequence 99% similar to it) coding without interruption for a human			900
XX	KSE336 polypeptide, a serine protein kinase, comprising the KSE336-1 and			960
XX	KSE336-2 splice variants appearing as ABG72382 and ABG72383. Also			1245
XX	included is a method of identifying an agent that modulates the			1020
XX	expression of KSE336 in brain, pancreas, brain progenitor or pancreas			1305
XX	progenitor cells comprising: (a) contacting a cell population comprising			1080
XX	the cells with a test agent under conditions effective for the test agent			1365
XX	to modulate the expression of KSE336; and (b) determining if the test			1140
XX	agent modulates the expression of KSE336. The polynucleotides are useful			1425
XX	as molecular targets or drug targets, and for detecting, diagnosing,			1200
XX	staging, monitoring, prognosticating, preventing or treating diseases or			1380
XX	conditions relating to brain and pancreas, such as astrocytoma,			1060
XX	meningioma, pancreatic adenocarcinoma, insulin-dependent diabetes			1185
XX	mellitus 2, heliocid peripapillary choriorretinal degeneration, Beckwith-			960
XX	Wiedemann syndrome or congenital hyperinsulinism. The method and			1245
XX	polynucleotides are useful in research, diagnosis, drug discovery,			1020
XX	therapy, clinical medicine, forensic science and pathology.			1305
XX	The gene for KSE336 is located on chromosome 11p5.5-pter.			1080
XX	The present sequence encodes the KSE366 variant KSE366-1.			1365
XX	Sequence 2908 BP; 603 A; 948 C; 879 G; 478 T; 0 other;			1060
XX	Query Match 99.9%; Score 1925.4; DB 25; Length 2908;			1185
XX	Best Local Similarity 99.9%; Pred. No. 0;			960
XX	Matches 1826; Conservative 0; Mismatches 1; Indels 0; Gaps 0;			1245
XX	1 ATGAAGGTGGAGCGGAGATCGGATCCTGAAGCTCATTTAGACACCCACGTCCTAAAG 60			1020
XX	286 ATGAAGGTGGAGCGGAGATCGGATCCTGAAGCTCATTTAGACACCCACGTCCTAAAG 345			1380
XX	61 CTGCACGAGCTTTATGA AAAAATAATTTGTAACCTGGTGCTAGAACAGTGTCAAGT 120			1060
XX	346 CTGCACGAGCTTTATGA AAAAATAATTTGTAACCTGGTGCTAGAACAGTGTCAAGT 405			1185
XX	121 GGTGAGCTCTTCGACTACTGTTGTAAGAGGGAGGCTGACCCCTAAGAGGCTCGGAAG 180			960
XX	406 GGTGAGCTCTTCGACTACTGTTGTAAGAGGGAGGCTGACCCCTAAGAGGCTCGGAAG 465			1245
XX	181 TTCTTTCCGCGACATCATCTCTGCGCTGGACTTCTGCGACAGCCACTCCATATGCCAGG 240			1020
XX	466 TTCTTTCCGCGACATCATCTCTGCGCTGGACTTCTGCGACAGCCACTCCATATGCCAGG 525			1305
XX	241 GATCTGAAACCTGAAACCTCTCTGCTGGACGAGAGAACACATCTCCGACATCGCAGCTTT 300			1080

Db 1606 ACGCGGAGGAGATGTCACACCTGACACAGAGTGTCTCCAGAGCTGGCGAAGAGTCC 1665
Qy 1381 TGGTTTGGAACTTCATCAGCTGAGAGAGAGAGAGATCTTCGTGCTCATCAAGAC 1440
Db 1666 TGGTTTGGAACTTCATCAGCTGAGAGAGAGAGAGATCTTCGTGCTCATCAAGAC 1725
Qy 1441 AAACCTTGAGCTTCATCAGCTGACATCGTGACAGCTTCCTGTGATTCACAGTCTC 1500
Db 1726 AAACCTTGAGCTTCATCAGCTGACATCGTGACAGCTTCCTGTGATTCACAGTCTC 1785
Qy 1501 AGCCACAGCTCATCTCCCAACAGAGCTTCGGGCGGAGTCAAGGCCACACGGGGGGCCA 1560
Db 1786 AGCCACAGCTCATCTCCCAACAGAGCTTCGGGCGGAGTCAAGGCCACACGGGGGGCCA 1845
Qy 1561 GCGGTGTTCCAGAACCGGTCAAGTTCAGAGTTGATATCACTACACGAGGGGTGGGAG 1620
Db 1846 GCGGTGTTCCAGAACCGGTCAAGTTCAGAGTTGATATCACTACACGAGGGGTGGGAG 1905
Qy 1621 GCGGAGAGGAGAACGGATCTACTCCGTACCTTCACCTGCTTCAAGGCCACAGCCGT 1680
Db 1906 GCGGAGAGGAGAACGGATCTACTCCGTACCTTCACCTGCTTCAAGGCCACAGCCGT 1965
Qy 1681 CGCTTCAAGAGGTGTGTGAGACCATCAGAGCCAGCTGTGAGCACACAGCCCGCT 1740
Db 1966 CGCTTCAAGAGGTGTGTGAGACCATCAGAGCCAGCTGTGAGCACACAGCCCGCT 2025
Qy 1741 GCGGCCACAGCTTGTGACACACACTAACTGTATGGAATGATGACGGGGCGGCTTTC 1800
Db 2026 GCGGCCACAGCTTGTGACACACACTAACTGTATGGAATGATGACGGGGCGGCTTTC 2085
Qy 1801 AAATGTGAATATCCCGAAAGTTAA 1827
Db 2086 AAATGTGAATATCCCGAAAGTTAA 2112

RESULT 5
ABX13584

ID ABX13584 standard; cDNA; 3364 BP.

XX AC ABX13584;

XX DT 10-FEB-2003 (first entry)

XX DE Human cDNA encoding serine protein kinase KSE336-2.

XX KW Human; ss; gene; chromosome 11p15.5-pter; astrocytoma; meningioma;
XX KW pancreatic adenocarcinoma; insulin-dependent diabetes mellitus 2;
XX KW helicoid peripapillary chorioretinal degeneration; brain; pancreas;
XX KW Beckwith-Wiedemann syndrome; congenital hyperinsulinism; KSE336-2.

XX OS Homo sapiens.

XX FH Key Location/Qualifiers

XX FT CDS 482..2239

XX FT /*tag= a

XX FT /product= "KSE336-2"

XX PN US6455292-B1.

XX PD 24-SEP-2002.

XX PF 16-AUG-2001; 2001US-0930181.

XX PR 16-AUG-2001; 2001US-0930181.

XX PA (ORIG-) ORIGENE TECHNOLOGIES INC.

XX PI Shu Y, Fan W, Kovacs KP, Zidanic M, Jay G;

XX DR WPI; 2003-066233/06.

XX DR P-PSDB; ABG72383.

XX PT New isolated polynucleotide coding without interruption for a human

KSE336 polypeptide useful for preventing or treating
diseases/conditions relating to brain and pancreas, e.g. meningioma,
insulin-dependent diabetes mellitus 2

Claim 4; Column 43-48; 34pp; English.

The invention relates to an isolated polynucleotide (its complement or a sequence 9% similar to it) coding without interruption for a human KSE336 polypeptide, a serine protein kinase, comprising the KSE336-1 and KSE336-2 splice variants appearing as ABG72382 and ABG72383. Also included is a method of identifying an agent that modulates the expression of KSE336 in brain, pancreas, brain progenitor or pancreas progenitor cells comprising: (a) contacting a cell population comprising the cells with a test agent under conditions effective for the test agent to modulate the expression of KSE336; and (b) determining if the test agent modulates the expression of KSE336. The polynucleotides are useful as molecular targets or drug targets, and for detecting, diagnosing or staging, monitoring, prognosticating, preventing or treating diseases or conditions relating to brain and pancreas, such as astrocytoma, meningioma, pancreatic adenocarcinoma, insulin-dependent diabetes mellitus 2, helicoid peripapillary chorioretinal degeneration, Beckwith-Wiedemann syndrome or congenital hyperinsulinism. The method and polynucleotides are useful in research, diagnosis, drug discovery, therapy, clinical medicine, forensic science and pathology. The gene for KSE336 is located on chromosome 11p15.5-pter. The present sequence encodes the KSE366 variant KSE366-2.

SQ Sequence 3364 BP; 639 A; 1078 C; 1069 G; 578 T; 0 other;

Query Match 99.2%; Score 1813; DB 25; Length 3364;

Best Local Similarity 99.8%; Pred. No. 0;

Matches 1827; Conservative 0; Mismatches 0; Indels 4; Gaps 1;

Qy 1 ATGAAGGTGAGCGGAGATCGCGATCTCTGAAGCTATTGAGCACCCCGCTTAAG 60
Db 409 ATGAAGGTGAGCGGAGATCGCGATCTCTGAAGCTATTGAGCACCCCGCTTAAG 468
Qy 61 CTGCACGACGTTTTATGAAAAACAAAAATATTT---GTACCTGTGTGTAGAACACGTGTC 116
Db 469 CTGCACGACGTTTTATGAAAAACAAAAATATTTGTAGTACTGTGTGTAGAACACGTGTC 528
Qy 117 AGGTGTGAGCTCTTCGACTACCTGTGTGAAGAGGGGAGCTGACCGCTAAGAGGCTCG 176
Db 529 AGGTGTGAGCTCTTCGACTACCTGTGTGAAGAGGGGAGCTGACCGCTAAGAGGCTCG 588
Qy 177 GAAGTTCTCCGCGAGATCATCTCGCTGGCTGGACTTCTGCGCAGAGCCACTTATGCCA 236
Db 589 GAAGTTCTCCGCGAGATCATCTCGCTGGCTGGACTTCTGCGCAGAGCCACTTATGCCA 648
Qy 237 CAGGATCTGAAACCTGAAACCTCTCTGTGGAGAGAGAAACAAACATCCGCATCGCAGA 296
Db 649 CAGGATCTGAAACCTGAAACCTCTCTGTGGAGAGAGAAACAAACATCCGCATCGCAGA 708
Qy 297 CTTTGGCATCGGCTCCCTGAGGTTGGCGACAGCTTTGGAGACAGCTGTGGGTCCCC 356
Db 709 CTTTGGCATCGGCTCCCTGAGGTTGGCGACAGCTTTGGAGACAGCTGTGGGTCCCC 768
Qy 357 CCACTACGCTGCCCGGAGGTGATCCGGGGGAGAGTATGACGGCCGGAGGCGGACGT 416
Db 769 CCACTACGCTGCCCGGAGGTGATCCGGGGGAGAGTATGACGGCCGGAGGCGGACGT 828
Qy 417 GTGAGCTGCGGGGTCTATCTCTGTGTGTGGGGGCTCTGCGCTTCGACGATGA 476
Db 829 GTGAGCTGCGGGGTCTATCTCTGTGTGTGGGGGCTCTGCGCTTCGACGATGA 888
Qy 477 CAACTTGGACAGCTGTGGAGAGGTGAACGGGGGCTGTTTCCATGCGCGACTTTAT 536
Db 889 CAACTTGGACAGCTGTGGAGAGGTGAACGGGGGCTGTTTCCACATGCGCGACTTTAT 948
Qy 537 CCGCCCGGACTGCCAGAGTCTGTACGGGGCATGATCGAGGTGGACCGCCGCGCT 596
Db 949 CCGCCCGGACTGCCAGAGTCTGTACGGGGCATGATCGAGGTGGACCGCCGCGCT 1008

QY 597 CAGCTAGACACATTCAGAAACACATATGGTATATAGGGGGCAAGATGAGCCGAACC 656
Db 1009 CAGCTAGACACATTCAGAAACACATATGGTATATAGGGGGCAAGATGAGCCGAACC 1068
QY 657 AGAGCAGCCCATTCCTCGCAAGGTGAGATCCGCTCGCTGCCAGCCTCGAGGACATCGA 716
Db 1069 AGAGCAGCCCATTCCTCGCAAGGTGAGATCCGCTCGCTGCCAGCCTCGAGGACATCGA 1128
QY 717 CCCCAGCGTCTGAGACGATGACATCTAGCTGCTGCTCCGAGACGGCAACAGCTGCT 776
Db 1129 CCCCAGCGTCTGAGACGATGACATCTAGCTGCTGCTCCGAGACGGCAACAGCTGCT 1188
QY 777 GCAGGACCTGCTGCTCCGAGGAGGAGAACAGGAGAGATGATTTATCTCTCTCTCTGA 836
Db 1189 GCAGGACCTGCTGCTCCGAGGAGGAGAACAGGAGAGATGATTTATCTCTCTCTCTGA 1248
QY 837 CCGGAAGAAAGGTATCCGAGCCAGGAGGATGAGACCTGCCCCCCCGGAACGAGATAGA 896
Db 1249 CCGGAAGAAAGGTATCCGAGCCAGGAGGATGAGACCTGCCCCCCCGGAACGAGATAGA 1308
QY 897 CCCTCCCGGAAGGTGAGTCCCGGATGCTGACCCGACGCGCAAGCGGGCCGAG 956
Db 1309 CCCTCCCGGAAGGTGAGTCCCGGATGCTGACCCGACGCGCAAGCGGGCCGAG 1368
QY 957 ACGCAATCCATGAGGTGCTCAGCGTGACGCGCGGCTCCCGGTGCTCGCGGGCG 1016
Db 1369 ACGCAATCCATGAGGTGCTCAGCGTGACGCGCGGCTCCCGGTGCTCGCGGGCG 1428
QY 1017 GGCCATTGAGATGCCCAGCAGCGCCAGAGGTCTCGGTCCATCAGCGGTGCTCTCAGG 1076
Db 1429 GGCCATTGAGATGCCCAGCAGCGCCAGAGGTCTCGGTCCATCAGCGGTGCTCTCAGG 1488
QY 1077 CTTTTCACGACCCACTCAGCAGCCCGCGGTGACCCCTCACCCCTCACAGGGGCGAG 1136
Db 1489 CTTTTCACGACCCACTCAGCAGCCCGCGGTGACCCCTCACCCCTCACAGGGGCGAG 1548
QY 1137 TCCCTCCCGCAAGGGGACACCTGTCCACACGCCAAAGAGAGCCCGGCTGGCAC 1196
Db 1549 TCCCTCCCGCAAGGGGACACCTGTCCACACGCCAAAGAGAGCCCGGCTGGCAC 1608
QY 1197 GCCCAACCCACGCCCGCTCAGCGCCAGGTGCGAGGGGTGCTCGAGGGCGGGCT 1256
Db 1609 GCCCAACCCACGCCCGCTCAGCGCCAGGTGCGAGGGGTGCTCGAGGGCGGGCT 1668
QY 1257 CACTCCATCAAGAACAGCTTCTGGGCTCACCCGCTTCCACGCCGGAACCTGCAAGT 1316
Db 1669 CACTCCATCAAGAACAGCTTCTGGGCTCACCCGCTTCCACGCCGGAACCTGCAAGT 1728
QY 1317 TCCGACGCCGAGGAGATGTCACACCTGACACACAGAGTCTGTCGCCAGAGCTGGCGAAGAA 1376
Db 1729 TCCGACGCCGAGGAGATGTCACACCTGACACACAGAGTCTGTCGCCAGAGCTGGCGAAGAA 1788
QY 1377 GTCTGTTTGGAACTTCATCAGCCTGAGAGAGGAGGAGAGATCTTCGTGTCATCAA 1436
Db 1789 GTCTGTTTGGAACTTCATCAGCCTGAGAGAGGAGGAGAGATCTTCGTGTCATCAA 1848
QY 1437 AGACAACTCTGAGCTCCATCAAGGTGACATCGTCAGCCTTCTCTGTCGATTCGCCAG 1496
Db 1849 AGACAACTCTGAGCTCCATCAAGGTGACATCGTCGACGCTTCTCTGTCGATTCGCCAG 1908
QY 1497 TCTCAGCCACAGCCTCATCTCCAAACAGGCTTCCGGGCGAGTACAGGCGCAGGGGGG 1556
Db 1909 TCTCAGCCACAGCCTCATCTCCAAACAGGCTTCCGGGCGAGTACAGGCGCAGGGGGG 1968
QY 1557 GCCAGCGTGTTCAGAAAGCGGTCAAGTTCAGGTTGATATCACTACAGGAGGTGG 1616
Db 1969 GCCAGCGTGTTCAGAAAGCGGTCAAGTTCAGGTTGATATCACTACAGGAGGTGG 2028
QY 1617 GGAGCGCAGAGGAGAACGCACTTACTCGTCACTCTCACTCTCTCAGGCCCCAG 1676
Db 2029 GGAGCGCAGAGGAGAACGCACTTACTCGTCACTCTCTCAGGCCCCAG 2088
QY 1677 CCGTCCGTTCAAGAGGGTGGTGGAGACCATCCAGGCCCGAGCTGTGAGCACACAGCACCC 1736

Db 2089 CCGTCGCTTCAAGAGGTGGTGGAGCCATCCAGGCCAGCTGCTGAGCACACAGCACCC 2148
QY 1737 GCCTGCGGCCAGCAGCTTGTGAGACACCACTAACTGTATGGAATGATGACGGGGCGCT 1796
Db 2149 GCCTGCGGCCAGCAGCTTGTGAGACACCACTAACTGTATGGAATGATGACGGGGCGCT 2208
QY 1797 TTCCAAATGTGGAATTATCCCGAAAGTTAA 1827
Db 2209 TTCCAAATGTGGAATTATCCCGAAAGTTAA 2239
RESULT 6
ABA02995
ID ABA02995 standard; cDNA; 2025 BP.
XX
AC ABA02995;
XX
DT 19-FEB-2002 (first entry)
XX
DE Human protein kinase 2246 coding sequence SEQ ID NO 3.
XX
KW Human; protein kinase 2246; cytostatic; immunomodulator; carcinoma;
anti-inflammatory; analgesic; cardiovascular; cancer; sarcoma;
cellular proliferation disorder; cellular differentiation disorder;
metastatic; haematopoietic disorder; leukaemia; immune disorder;
inflammatory disorder; arthritis; autoimmune disease; diabetes mellitus;
psoriasis; Crohn's disease; cardiovascular disease; virus; pain;
gene therapy; ss.
XX
OS Homo sapiens.
XX
PH Key Location/Qualifiers
FT CDS 1..2025
FT /tag= a
FT /product= "protein kinase 2246"
XX
XX WO200181588-A2.
XX
XX 01-NOV-2001.
XX
XX 25-APR-2001; 2001WO-US13784.
XX
XX 25-APR-2000; 2000US-199391P.
XX
XX (WILL-) MILLENNIUM PHARM INC.
XX
XX Meyers R;
XX
XX WPI; 2002-049281/06.
XX
XX P-PSDB; AAM47830.
XX
XX New protein kinase nucleic acid and polypeptide molecules, designated
PT 2246, useful for diagnosing, preventing or treating cancer or a
PT cellular proliferation/differentiation disorders, e.g. carcinoma,
PT sarcoma or leukaemias
XX
XX Claim 1; Fig 1; 111pp; English.
XX
XX The invention relates to the human protein kinase 2246 gene and the
CC the isolated encoded polypeptide with cytostatic, immunomodulator,
CC anti-inflammatory, analgesic and cardiovascular activity. The 2246
CC nucleic acid and polypeptide are useful for diagnosing, preventing or
CC treating a subject having cancer or a cellular proliferation and/or
CC differentiation disorder or at risk of developing cancer or a cellular
CC proliferation and/or differentiation disorder. In particular, the
CC disorder includes carcinoma, sarcoma, metastatic or haematopoietic
CC disorders (e.g. leukaemias) or cancers of the lung, breast, thyroid, head
CC neck, prostate or genito-urinary tract. The 2246 nucleic acid and
CC polypeptide are also useful for treating immune disorders, e.g.
CC inflammatory (e.g. respiratory inflammation or arthritis), autoimmune
CC disease (e.g. diabetes mellitus, psoriasis, Wegener's granulomatosis,
CC Crohn's disease or Grave's disease), for treating cardiovascular

CC	diseases, endothelial cell disorder, viral diseases or pain. The nucleic acid and polypeptide are also useful for evaluating the efficacy of a treatment of cancer or a cellular proliferation and/or differentiation disorder. The nucleic acid is useful for gene therapy. The present sequence is that of the 2246 coding sequence.
XX	
SQ	Sequence 2025 BP; 442 A; 641 C; 616 G; 326 T; 0 other;
	Query Match Best Local Similarity 96.3%; Score 1759.6; DB 24; Length 2025; Matches 1762; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
Qy	1 ATGAAGGTGGACGCGGAGATCGCATCTCTGAAGCTCATTGAGCACCCCACGTCGTCTAAAG 60
Db	181 ATGAAGGTGGAGCGGGAGATCGCATCTCTGAAGCTCATTGAGCACCCCACGTCGTCTAAAG 240
Qy	61 CTGCACGACGTTTATGAAGAACAATAATTTGTACTGTGTCTAGACACGTCGTCAAGT 120
Db	241 CTGCACGACGTTTATGAAGAACAATAATTTGTACTGTGTGTCTAGAACACGTCGTCAAGT 300
Qy	121 GGTGAGCTCTTCGACTACCTCTGCTGAAGAAGGCGGCTGACGCCCTTAAGGAGGCTCGGAAG 180
Db	301 GGTGAGCTCTTCGACTACCTCTGCTGAAGAAGGCGGAGCTGACGCCCTTAAGGAGGCTCGGAAG 360
Qy	181 TTCCTTCGGCGAGATCATCTCTGCGCTGACCTCTTGCCACAGCCACTCCATATGCCACAGG 240
Db	361 TTCCTTCGGCGAGATCATCTCTGCGCTGAGCTCTTGCCACAGCCACTCCATATGCCACAGG 420
Qy	241 GATCTGAACCTGTAAAACCTCCTCTGTGACACGAGAGAACATATCCGCAATCCGACACTTT 300
Db	421 GATCTGAACCTGTAAAACCTCCTCTGTGACACGAGAGAACATATCCGCAATCCGACACTTT 480
Qy	301 GGCAATGGCGTCCCTCGAGGTTGGCGACAGCCCTGTTGGAGACCAGCTGTGGGTCCCCCCAC 360
Db	481 GGCAATGGCGTCCCTCGAGGTTGGCGACAGCCCTGTTGGAGACCAGCTGTGGGTCCCCCCAC 540
Qy	361 TACGCTGCCCCGAGGTGATCCGGGGGAGAAATGATGACGCCCGAAGCGGACGTGTGG 420
Db	541 TACGCTGCCCCGAGGTGATCCGGGGGAGAAATGATGACGCCCGGAGCGGACGTGTGG 600
Qy	421 AGCTGCGCGCTCATCTGTTGCGCTTGCTGTGTGGGGCTCTGCCCTTCGACGATGACAC 480
Db	601 AGCTGCGCGCTCATCTGTTGCGCTTGCTGTGTGGGGCTCTGCCCTTCGACGATGACAC 660
Qy	481 TTGCGACAGCTGCTGGAGAAAGGTCAAAGCGGGCGCTGTTCCAATATGCGCGCACTTTATCCG 540
Db	661 TTGCGACAGCTGCTGGAGAAAGGTCAAAGCGGGCGCTGTTCCAATATGCGCGCACTTTATCCG 720
Qy	541 CCCGACTGCCAGAGTCTGCTACGGGGCATGATCGAGGTGGACGCCGCA CGCGGCTCAAG 600
Db	721 CCCGACTGCCAGAGTCTGCTACGGGGCATGATCGAGGTGGACGCCGCA CGCGGCTCAAG 780
Qy	601 CTAGAGCACATTAGAAACACATATGTTATATAGGGGCAAGATGAGCCCGCAACCTAGAG 660
Db	781 CTAGAGCACATTAGAAACACATATGTTATATAGGGGCAAGATGAGCCCGCAACCTAGAG 840
Qy	661 CAGCCCAATTCTCGCAAGGTTCAGATCCGCTCGCTGCCCAGCCTTGGAGGACATCGACCCC 720
Db	841 CAGCCCAATTCTCGCAAGGTTCAGATCCGCTCGCTGCCCAGCCTTGGAGGACATCGACCCC 900
Qy	721 GACGTGTGGACAGCATGCACTCACTGCGGTGCTTCGAGACCGCAACCAAGCTCTGCGAG 780
Db	901 GACGTGTGGACAGCATGCACTCACTGCGGTGCTTCGAGACCGCAACCAAGCTCTGCGAG 960
Qy	781 GACCTGTGTCGAGGAGGAGAACACAGAGAAATGATTTACTTCTCTCTCTGGACCGG 840
Db	961 GACCTGTGTCGAGGAGGAGAACACAGAGAAATGATTTACTTCTCTCTCTGGACCGG 1020
Qy	841 AAAGAAAGGTACCCGAGCCAGGAGATGAGGACCTGCCCCCGCGMAACGAGATAGACCT 900
Db	1021 AAAGAAAGGTACCCGAGCCAGGAGATGAGGACCTGCCCCCGCGMAACGAGATAGACCT 1080
Qy	901 CCCCGAAGCGTGTGGACTCCCGCATGCTGAACCGGCAACGCAAGCGCGCGCCAGAACGC 960

KW metastatic; haematopoietic disorder; leukaemia; immune disorder;
KW inflammatory disorder; arthritis; autoimmune disease; diabetes mellitus;
KW psoriasis; Crohn's disease; cardiovascular disease; virus; pain;
XX gene therapy; 88.
XX Homo sapiens.
XX OS
XX FH Location/Qualifiers
XX CDS 51..2075
XX FT /*tag= a
XX FT /product= "protein kinase 2246"
XX PN WO200181588-A2.
XX XX
XX PD 01-NOV-2001.
XX XX
XX PF 25-APR-2001; 2001WO-US13784.
XX XX
XX PR 25-APR-2000; 2000US-199391P.
XX XX (MILL-) MILLENNIUM PHARM INC.
XX PA
XX XX
XX PI Meyers R;
XX XX
XX DR WPI; 2002-049281/06.
XX DR P-PSDB; AAM47830.
XX XX
XX PT New protein kinase nucleic acid and polypeptide molecules, designated
XX FT 2246, useful for diagnosing, preventing or treating cancer or a
XX FT cellular proliferation/differentiation disorders, e.g. carcinoma,
XX PT sarcoma or leukaemias
XX XX
XX PS Claim 1; Fig 1; 11pp; English.
XX CC
XX CC The invention relates to the human protein kinase 2246 gene and the
XX CC the isolated encoded polypeptide with cytostatic, immunomodulator,
XX CC anti-inflammatory, analgesic and cardiovascular activity. The 2246
XX CC nucleic acid and polypeptide are useful for diagnosing, preventing or
XX CC treating a subject having cancer or a cellular proliferation and/or
XX CC differentiation disorder or at risk of developing cancer or a cellular
XX CC proliferation and/or differentiation disorder. In particular, the
XX CC disorder includes carcinoma, sarcoma, metastatic or haematopoietic
XX CC disorders (e.g. leukaemias) or cancers of the lung, breast, thyroid, head
XX CC neck, prostate or genito-urinary tract. The 2246 nucleic acid and
XX CC polypeptide are also useful for treating immune disorders, e.g.
XX CC inflammatory (e.g. respiratory inflammation or arthritis), autoimmune
XX CC disease (e.g. diabetes mellitus, psoriasis, Wegener's granulomatosis,
XX CC Crohn's disease or Grave's disease), for treating cardiovascular
XX CC diseases, endothelial cell disorder, viral diseases or pain. The nucleic
XX CC acid and polypeptide are also useful for evaluating the efficacy of a
XX CC treatment of cancer or a cellular proliferation and/or differentiation
XX CC disorder. The nucleic acid is useful for gene therapy. The present
XX CC sequence is that of the 2246 encoding cDNA.
XX XX
SQ Sequence 2217 BP; 476 A; 715 C; 682 G; 344 T; 0 other;
Query Match 96.3%; Score 1759.6; DB 24; Length 2217;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 1762; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
Qy 1 ATGAAGGTGGAGCGGAGATCGGATCCTGAAGCTCATTTAGACACCCCGCTCTAAAG 60
Db 231 ATGAAGGTGGAGCGGAGATCGGATCCTGAAGCTCATTTAGACACCCCGCTCTAAAG 290
Qy 61 CTGACGACGCTTTATGAACAAAAATATTTGACTGTGTGTAGAACACGCTGTAGGT 120
Db 291 CTGACGACGCTTTATGAACAAAAATATTTGACTGTGTGTAGAACACGCTGTAGGT 350
Qy 121 GGTGAGCTCTTCGACTACTCTGCTGAAGAGGGAGGCTGAGCCCTAAGGAGGCTCGGAAG 180
Db 351 GGTGAGCTCTTCGACTACTCTGCTGAAGAGGGAGGCTGAGCCCTAAGGAGGCTCGGAAG 410
Qy 181 TTCTTCGGGCAGATCATCTCTGCGCTGGACTTTCTGCCACAGCCACTCCATATATGCCACAGG 240

Db 411 TTCTTCGGGCAGATCATCTCTGCGCTGGACTTTCTGCCACAGCCACTCCATATGCCACAGG 470
Qy 241 GATCTGAACCTGAAACCTCTCTGTGGACAGAGAAACAAATCCGATCCGATCCGACATTT 300
Db 471 GATCTGAACCTGAAACCTCTCTGTGGACAGAGAAACAAATCCGATCCGACATTT 530
Qy 301 GGCATGGCGTCCCTGCAGGTTGGCCGACAGCCCTTTGGAGACAGCTGTGGGTCCGCCAC 360
Db 531 GGCATGGCGTCCCTGCAGGTTGGCCGACAGCCCTTTGGAGACAGCTGTGGGTCCGCCAC 590
Qy 361 TAGCGCTGCCCGGAGGTGATCCGGGGGAGAAATATGACGGCCGGAAGCGGACGTGTGG 420
Db 591 TAGCGCTGCCCGGAGGTGATCCGGGGGAGAAATATGACGGCCGGAAGCGGACGTGTGG 650
Qy 421 AGCTGGCGGCTCATCTCTGTTCGCTGTGGGTGGGGCTCTGCCCTTCGACATGACAAAC 480
Db 651 AGCTGGCGGCTCATCTCTGTTCGCTGTGGGTGGGGCTCTGCCCTTCGACATGACAAAC 710
Qy 481 TTGCGACAGCTGCTGGAGAGGTGAAGCGGGCGCTGTTCCACATGCGCGACTTTATCCCG 540
Db 711 TTGCGACAGCTGCTGGAGAGGTGAAGCGGGCGCTGTTCCACATGCGCGACTTTATCCCG 770
Qy 541 CCCGACTGCCAGAGTCTGCTACGGGCGATGATCGAGGTGGACCGCGCACCGCCCTCACG 600
Db 771 CCCGACTGCCAGAGTCTGCTACGGGCGATGATGAGGTGGACCGCGCACCGCCCTCACG 830
Qy 601 CTAGAGCACAATTCAGAAACACATATGATATAGGGGGCAAGAAAGAGCCCGCAACAGAG 560
Db 831 CTAGAGCACAATTCAGAAACACATATGATATAGGGGGCAAGAAAGAGCCCGCAACAGAG 890
Qy 661 CAGCCCTCTTCGCAAGGTGCAGATCCGCTCGCTGCCAGCTGGAGGACATCGACCC 720
Db 891 CAGCCCTCTTCGCAAGGTGCAGATCCGCTCGCTGCCAGCTGGAGGACATCGACCC 950
Qy 721 GACGTGCTGGACAGCATGCACTCACTGGGCTGCTTCCGAGACCGCAACAGCTGTGCGAG 780
Db 951 GACGTGCTGGACAGCATGCACTCACTGGGCTGCTTCCGAGACCGCAACAGCTGTGCGAG 1010
Qy 781 GACCTGCTGTCGAGGAGGAGAACAGGAGAGATGATTTACTTCTCTCTCTGGACCGG 840
Db 1011 GACCTGCTGTCGAGGAGGAGAACAGGAGAGATGATTTACTTCTCTCTCTGGACCGG 1070
Qy 841 AAAGAAAGTACCCGAGCCAGAGGATGAGGACCTGCCCCCGGAAACAGATAGACCT 900
Db 1071 AAAGAAAGTACCCGAGCCAGAGGATGAGGACCTGCCCCCGGAAACAGATAGACCT 1130
Qy 901 CCCGGAAGCGTGTGGACTCCCGGATGCTGAAACCGGACGCAAGCGGCGGCGCAGAACGC 960
Db 1131 CCCGGAAGCGTGTGGACTCCCGGATGCTGAAACCGGACGCAAGCGGCGGCGCAGAACGC 1190
Qy 961 AAATCATGAGGTGCTCAGCTGACGAGGGGGCTCCCGGTCCTCGGCGGGCGG 1020
Db 1191 AAATCATGAGGTGCTCAGCTGACGAGGGGGCTCCCGGTCCTCGGCGGGCGG 1250
Qy 1021 ATTGAGATGGCCAGCAGCGCCAGAGGTCTCGGTCCATCAGCGGTGCTCTCAGCGCTT 1080
Db 1251 ATTGAGATGGCCAGCAGCGCCAGAGGTCTCGGTCCATCAGCGGTGCTCTCAGCGCTT 1310
Qy 1081 TCCACAGCCCACTCAGCAGCCCGGGGTGACCCCTCACCCCTCAACAAAGGGGAGTCC 1140
Db 1311 TCCACAGCCCACTCAGCAGCCCGGGGTGACCCCTCACCCCTCAACAAAGGGGAGTCC 1370
Qy 1141 CTCGCCACCCCAAGGGGACAGCTGTCACGCGCAAGAGAGCCCGGCTGGCAGGCC 1200
Db 1371 CTCGCCACCCCAAGGGGACAGCTGTCACGCGCAAGAGAGCCCGGCTGGCAGGCC 1430
Qy 1201 AACCCCAACGCCCGCTCCAGCGCCAGCGTCGAGGGGTGCTCCCTGGAGGGCGGCTCAAC 1260
Db 1431 AACCCCAACGCCCGCTCCAGCGCCAGCGTCGAGGGGTGCTCCCTGGAGGGCGGCTCAAC 1490
Qy 1261 TCCATCAAGAACAGCTTTCTGSGCTCACCCCGTTTCACCCCGGAAACTGCAAGTTCCG 1320

Db 1491 TCCATCAAGAACAGCTTCTTGGGCTCACCCCGCTTCCACCCCGGAAACTGCAAGTTCCG 1550
 QY 1321 AGCGGAGGAGATGTCACAACTGACACAGAGTCTGTCCTCCAGAGCTGGCGAAGATGCC 1380
 Db 1551 AGCGGAGGAGATGTCACAACTGACACAGAGTCTGTCCTCCAGAGCTGGCGAAGATGCC 1610
 QY 1381 TGGTTTGGAACTTCATCAGCTGGAGAGGAGGAGATCTTCGTGGTTCATCAAGAC 1440
 Db 1611 TGGTTTGGAACTTCATCAGCTGGAGAGGAGGAGATCTTCGTGGTTCATCAAGAC 1670
 QY 1441 AAACCTCTGAGCTCCATCAAGGCTGACATCGTCAGCGCTTCTCTGTGATTCACAGTCTC 1500
 Db 1671 AAACCTCTGAGCTCCATCAAGGCTGACATCGTCAGCGCTTCTCTGTGATTCACAGTCTC 1730
 QY 1501 AGCCACAGCTCATCTCCCAACAGAGCTTCCGGGCCGAGTACAAGGCCACCGGGGGGCCA 1560
 Db 1731 AGCCACAGCTCATCTCCCAACAGAGCTTCCGGGCCGAGTACAAGGCCACCGGGGGGCCA 1790
 QY 1561 GCCGTGTTCCAGAGCCGGTCAAGTTCCAGGTTGATATCACTTACACGAGGGTGGGGAG 1620
 Db 1791 GCCGTGTTCCAGAGCCGGTCAAGTTCCAGGTTGATATCACTTACACGAGGGTGGGGAG 1850
 QY 1621 GCGCAGAAGGAGAACGGCATCTACTCCGTCACTTCCAGCTTCCAGCTTCCAGGCCCCAGCGT 1680
 Db 1851 GCGCAGAAGGAGAACGGCATCTACTCCGTCACTTCCAGCTTCCAGCTTCCAGGCCCCAGCGT 1910
 QY 1681 CGCTTCAAGAGGGTGGAGACCACTCAGGCCAGCTGCTGAGCACAACAGACCCGCT 1740
 Db 1911 CGCTTCAAGAGGGTGGAGACCACTCAGGCCAGCTGCTGAGCACAACAGACCCGCT 1970
 QY 1741 GCGGCCAGAGCTTGTGAGACCACTC 1766
 Db 1971 GCGGCCAGAGCTTGTGAGAACCCCC 1996

RESULT 8

AAL51889

ID AAL51889 standard; DNA; 2799 BP.

XX AC AAL51889;

XX DT 08-MAY-2003 (first entry)

XX XX Human cell cycle-regulatory factor Cdr2 coding sequence.

DE Human; cell cycle-regulatory factor; Cdr2; kinase; proliferative disease;

XX Human; cell cycle-regulatory factor; Cdr2; kinase; proliferative disease;

XX anticancer agent; wound-healing drug; gene; ds.

XX OS Homo sapiens.

XX Key Location/Qualifiers

FH 50..2314

FT CDS /*tag= a

FT FT /product= "Human Cdr2 protein"

XX WO200299110-A1.

XX 12-DEC-2002.

XX 03-JUN-2002; 2002WO-JP05411.

XX 04-JUN-2001; 2001JP-0168792.

XX (TAIH) TAIHO PHARM CO LTD.

XX (NAKA/) NAKANISHI M.

XX Nakanishi M;

XX WPI; 2003-156857/15.

XX P-PSDB; AA016604.

XX Cell cycle-regulatory factor Cdr2 with kinase activity and encoded

XX Gene, applicable in diagnosis of and screening drugs for proliferative

PT

PT diseases e.g. anticancer agents and wound-healing drugs -

XX Claim 2; Fig 2 A-C; 63pp; Japanese.

XX The invention comprises the amino acid and coding sequence of the human

CC cell cycle-regulatory factor Cdr2 (with kinase activity). The DNA and

CC protein sequences of the invention are useful in diagnosing and screening

CC drugs for proliferative diseases (e.g. anticancer agents and wound-

CC healing drugs). The present DNA sequence encodes the human cell cycle-

CC regulatory factor Cdr2 protein.

XX Sequence 2799 BP; 570 A; 908 C; 831 G; 490 T; 0 other;

SQ

Query Match 49.5%; Score 904.2; DB 25; Length 2799;

Best Local Similarity 70.9%; Pred. No. 4.6e-188;

Matches 1344; Conservative 0; Mismatches 428; Indels 123; Gaps 6;

QY 1 ATGAAGTGGAGCGGAGATCGCGATCTCTGAAGCTCATTAAGCACCCCAAGCTCTAAAG 60

Db 275 ATGAAGTGGAGCGGAGATCGCGATCTCTGAAGCTCATTAAGCACCCCAAGCTCTCAAG 334

QY 61 CTCACGACGCTTTATGAACAAACAAATATTTCTACCTGGTGTAGAACACGCTCTCAGGT 120

Db 335 CTCACGACGCTTTATGAACAAACAAATATTTCTACCTGGTGTAGAACACGCTCTCAGGT 394

QY 121 GGTGAGCTTTTCGACTACCTGGTGAAGAGGGAGGCTGACGCTTAAGGAGGCTTCGGAAG 180

Db 395 GGTGAGCTTTTCGACTACCTGGTGAAGAGGGAGGCTGACGCTTAAGGAGGCTTCGGAAG 454

QY 181 TTCTTCGGGAGATCATCTCTGCGCTGGACTTCTGCGACAGCCATCTCATATGCGACAGA 240

Db 455 TTCTTCGGGAGATCATCTCTGCGCTGGACTTCTGCGACAGCCATCTCATATGCGACAGA 514

QY 241 GATCTGAACCTGAACACCTCTGCTGCGAGAGAACAAACATCCGATCGGAGACTTT 300

Db 515 GACCTAAGCCCGAGAACCTCTGCTGCGAGAGAACAAACATCCGATCGGAGACTTT 574

QY 301 GGCATGGCGTCCCTGACAGGTTGGCGACAGCTGTTGGAGACCAAGCTTGGGTCCCCCAG 360

Db 575 GGCATGGCGTCCCTGACAGGTTGGCGACAGCTTGGGTCCCCCAGCTTGGGTCCCCCAG 634

QY 361 TAGCCTGCCCGAGGATCGCGGGGAGAAATATGACGCGCGGAAAGGCGAGCTGTGG 420

Db 635 TATGCGGTGTCAGAGGTGATTAAGGGGGGAAAAATATGATGGCGCGGCGGAGACATGG 694

QY 421 AGCTGCGGCTCATCTCTGCTGCTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 480

Db 695 AGCTGCGGCTCATCTCTGCTGCTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 754

QY 481 TTGCGACAGCTGCTGGAGAAAGGTGAAGCGGGGGTGTTCACATGCGCGACTTTATCCCG 540

Db 755 CTCGCCAGCTGCTGGAGAAAGGTGAAGCGGGGGTGTTCACATGCGCGACTTTATCCCG 814

QY 541 CCGAGCTGCCAGAGTCTGCTACGGGGCATGATGAGGTGACGCGCGACCGCGGCTCAGG 600

Db 815 CCAGATTGCCAGAGCTCTCTGAGGGGAAATGATGAGGTGAGGCGCCGAAAAAAGGCTCAGT 874

QY 601 CTAGAGCACATTCAGAAACACATATGATATAGGGGGGAGAAATGAGCGCGAGCC--- 656

Db 875 CTGAGGCAAAATTCAGAAACACATATGATATAGGGGGGAGAAATGAGCGCGAGCC--- 934

QY 657 --AGAGCAGGCCCATTCCTCGCAAGGTGCAGATCCGCTCGCTGCGCCAGCCTGGAGGACATC 714

Db 935 CTGAGGCCAGCCCTGCGCGCGGGTAGCCATCGGAGCCTGCCATCAACCGAGAGCTG 994

QY 715 GACCCGAGCTGCTGGAGACAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 774

Db 995 GACCCGAGCTGCTGAGAGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1054

QY 775 CTGAGGACCTGCTGCTGCGAGGAGGAGAACACAGAGAGATGATTTACTTCTCTCTCTG 834

Db 1055 CATCGCGAGCTGCGCAGTGGAGGAGAACCAAGAAAGATGATATATATCTGCTTTTG 1114

	D	b	G G C A T G G C G T C C T T G C A G G T G G G G A C A G C C T C T T G A G A C C A G T C G G G T C C C C C A T	360
301	Qy	T A C G C C T G C C C C G A G G T G A T C C G G G G G A G A A G T A T A C G C C G G A A G G O G G A C T G T G G	420	
361	D b	T A T G C G T C C A G A G T G A T T A G G G G G A A A T A T A T G T G C C G C C G G G A G A C A T G T G G	420	
421	Qy	A G T G C G G G T C A T C T T G T G C C C T T G T G G G G C T C T G C C C T T C G A C G T T T A T T C C G	480	
421	D b	A G C T G T G A G T C A T C C T T T C G C C C T G C T G T G G G G C T C T G C C C T T G A T G A C G A C A C	480	
481	Qy	T T G C G A C A G C T G T G G A R A G G T G A A G G G G G C G T G T T C A C A T G C C G C A C T T A T T C C G	540	
481	D b	C T C G C C A G C T G C T G S A A A G G T G A A A C G G G G C G T T T C C A C A T G C C C C A C T T C A T C C T	540	
541	Qy	C C G A C T G C C A G A G T C T G C T A C G G G G C A T G A T C A G G T G A C G C C G C A C C C C G C C T C A C G	600	
541	D b	C C A G A T T G C C A G A G C C T C C T G A G G G G A T G A T C G A A G T G A G C C G G A A A A A A G G C T C A G T	600	
601	Qy	C T A G A C A C A T T C A G A A C A C A T A T G T A T A T A G G G G G C A A G A T A G C C C G A A C C ---	656	
601	D b	C T G A G C A A A T T C A G A A C A C C T T G G T A C C T A G C G G G A A A C A C A G A G C C A G A C C C G T G C	660	
657	Qy	-- A G A C A G C C C A T T C C T C G C M A G G T G C A G A T C G C T G C T G C C C A G C C T G H A G E A C A T C	714	
661	D b	C T G A G C C A G C C C T G B C C G C C G G T A G C C A T G C G A G C C T G C C A T C C A A C G G A G A G C T G	720	
715	Qy	G A C C C G A C G T G T G A C A G C A T G C A C T C A C T G G G C T G T T C G A G A C C G C A A C A A G C T G	774	
721	D b	G A C C C G A C G T C C T A G A G A C A T G G C A T C A C T G G G C T G C T T A G G H A C C C G G A G A G C T G	780	
775	Qy	C T G C A G A C C T G T G T C C G A G G A G A A C C A G G A G A A G A T G A T T T A C T T C C T C C T C C T G	834	
781	D b	C A T C G C A G C T G C G A G T G A G G A G A A C C A A G A A A A G A T A T A T A T T A T C T G C T T T T G	840	
835	Qy	G A C C G A A A A A G G T A C C G A C C A G G A G A T G A G A C C T G C C C C C C C C C G G A A C G A G A T A	894	
841	D b	G A T C G A A G A G C G G T A T C C A G C T G T G A G A C C A G G A C C T G C C T C C C G G A A T G A T T	900	
895	Qy	G A C C C T C C C C G A A G C G T G T G G A C T C C C G A T G C T G A A C C G G A C G G A A G C G C G C C C A	954	
901	D b	G A C C C C C C C G A A A C U G T G B A T T C T C C A T G C T G A G C C G T C A C G G A A G C G C G A C C A	960	
955	Qy	G A A C G C A A A T C C A T G A G G T G C T C A G C G T G A C G A - - - - - C G C G C G C T C C C C G G T G	1005	
961	D b	G A G C G A A G T C C A T G G A A G T C C T G A G C A T C A C G A T G C C G G G G T G T G C T C C C C T G T A	1020	
1006	Qy	C T T G C G G G G G G C A T T G A T G C C C A G G A C G G C C A G A G T C T C G G T C C A T C A G C G G T	1065	
1021	D b	C C C A C C C G A C G G G C T T T G S A G A T G G C C C A G C A C C A G A G A T C C C G T A G C G T C A G T G A	1080	
1066	Qy	G C C T C C T C A G S C C T T T C A C C A G C C C A C T C A G A G C C C C G C G G - - - - -	1108	
1081	D b	G C C T C C A C G G T C T G T C C T C A G C C C T C T A A G A C C C C A A G A G T C G G T C T T T T C C T T T	1140	
1109	Qy	- - - - -	1108	
1141	D b	T C A C C G A G C C G G G G T G G A G A T A G G C T C G A G C G G G G C T C C C C G A C T T C C A A A A C G	1200	
1109	Qy	- - - - - T G A C C C C T C A C C C	1122	
1201	D b	C A G A C G T G C C T T C T C G G G G C C C A G G G T G G G G G C G C C G G G A G A C G C C C G C C C C C	1260	
1123	Qy	T C A C C A G G G G A G T C C C C T C C C C - - - - -	ACC 1149	
1261	D b	A G T G C C C G C T C C A C A C C C C T G C C C G G C C C C C C C A G C T C C C C G G C T C C T C T G G G G G A C C	1320	
1150	Qy	C C C A A G G G A C A C C T G T C C A C A C C C A A A G A G A G C C C G G C T G G C A C G C C C A A C C C C A C G	1209	
1321	D b	C C C T T G C A C T G C C T C T G C A C A C G C C C G G G C A G T C C C A C C G G A C C C C G G G G A C A A C A	1380	
1210	Qy	C C C C G T C C A G C C C C - - - - - A G C G T C G A G G G T G C C C T T G A G G G C G C G G C T C A A T C C	1363	

PT New human kinase proteins useful for diagnosis, drug screening, and
PT clinical trial monitoring, treatment of disorders and diseases, and
XX cosmetic and nutritional applications

PS Claim 5; Page 39; 47pp; English.

XX The present sequence is the coding sequence for a novel human kinase. The
CC genomic locus encoding the kinase is thought to be on human chromosome
CC 19. The kinase and its coding sequence are useful for diagnosis, drug
CC screening, clinical trial monitoring, treatment of disorders and
CC diseases, and cosmetic and nutritional applications.

XX SQ Sequence 2337 BP; 479 A; 786 C; 686 G; 386 T; 0 other;

Query Match 46.6%; Score 851.4; DB 25; Length 2337;
Best Local Similarity 68.9%; Pred. No. 1.6e-176;
Matches 1356; Conservative 0; Mismatches 416; Indels 195; Gaps 6;

QY 1 ATGAAGTGGAGGGGAGATCGCGATCCTGAAGCTCATTTGAGCAGCCCAACGTCCTTAAAG 60
DB |||||
QY 226 ATGAGGTGGAGGGGAGATCGCGATCCTGAAGCTCATTTGAGCAGCCCAACGTCCTCAAG 285
DB |||||
QY 61 CTCGACGACCTTTATGAAACAAATAATTTGTAAGCTGAGTGTGAGTGTGAGTGTGAGTGT 120
DB |||||
QY 286 CTCACGACGCTACGAGAACAGAAATATTTGTAAGCTGAGTGTGAGTGTGAGTGTGAGTGT 345
DB |||||
QY 121 GGTGAGCTCTTCGACTACCTGTTGAGAGGGGAGGCTGAGCCTTAAGGGCTCGGAG 180
DB |||||
QY 346 GGTGAGCTATTCGACTACCTGTTGAGAGGGGAGGCTGAGCCTTAAGGGCTCGGAG 405
DB |||||
QY 181 TTCTTCGGCAGATCATCTCTGGCTGGAATTTGTCACAGGCACTTCCATATGCCACAGG 240
DB |||||
QY 406 TTCTTCGGCAGATTTGTTCTGGCTGGAATTTGTCACAGGCTTCCATATGCCACAGG 465
DB |||||
QY 241 GATCTGAAACCTGTAACCTCTGCTGGAAGAGAAACAAATTCGATTCGAGACTTT 300
DB |||||
QY 466 GACCTAAAGCCGAGAACCTCTTTTGGATGAGAAACAAATTCGATTCGAGACTTT 525
DB |||||
QY 301 GGCATGCGCTCCCTGAGGTTGGGACAGCTGTTGGAGACCAAGCTGTGGGTCCTCCAC 360
DB |||||
QY 526 GGCATGCGCTCCCTGAGGTTGGGACAGCTTCTTGGAGACCAAGCTGTGGGTCCTCCAC 585
DB |||||
QY 361 TACGCTCCGCGAGGTGATCCGGGGGAGAAATATGACGCGCGGAGGCGAGCTGTGG 420
DB |||||
QY 586 TATGCGTGTCCAGAGGTGATTAAGGGGAAATATGATGCGCGCGGACAGATGTG 645
DB |||||
QY 421 AGCTGCGGCTGATCTCTGTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 480
DB |||||
QY 646 AGCTGCGGCTGATCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 705
DB |||||
QY 481 TTGCGACAGCTGTGGAGAGGTGAAAGCGGGGCTGTTCCACATGCGGCACTTTATCCCG 540
DB |||||
QY 706 CTGCGGACAGCTGTGGAGAGGTGAAAGCGGGGCTTCCACATGCGGCACTTTATCCCT 765
DB |||||
QY 541 CCCGACTGCCAGAGCTGTCTACGCGGGGATGATGAGGTGAGCGCGCACGCGGCTCACG 600
DB |||||
QY 766 CCAGATTGCCAGAGCTCTCTGAGGAGGATGATCGAAGTGGAGCCGAAAAAGGCTCAGT 825
DB |||||
QY 601 CTAGAGCATTTCAGAAACACATATGTTATATAGGGGGGAGAAATGAGCCCGAACCC 656
DB |||||
QY 826 CTGAGGCAATTCAGAAACATCTCTGTTGTTACCTAGCGGGGAAACACAGAGCCAGCCG 885
DB |||||
QY 657 --AGAGAGCCCATCTCTCGAAGTGCAGATCCGCTCGCTGCTGCTGCTGCTGCTGCTGCT 714
DB |||||
QY 886 CTGAGGAGCCAGCCCTGGCGCGGGGTAGCCATGCGGAGCTTGCCATCCAAAGGAGCTG 945
DB |||||
QY 715 GACCCGAGCTGTGGAAGCAGCATCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 774
DB |||||
QY 946 GACCCGAGCTCTAGAGAGCATGGCATCATCTGGGCTGCTTTCAGGGAGCCGCGAGGCTG 1005
DB |||||
QY 775 CTGAGGAGCTGTGCTGCGAGGAGAGAAACAGAGAGAGATGATTACTTCTCTCTCTG 834
DB |||||
QY 1006 CATCGGAGCTGCGAGTGGAGGAGAGAACCAAGAAAGATGATATATATCTGCTTTTG 1065
DB |||||

QY 1726 ACACACGACCGCGCTGCGGCCGACGACTTGTGACACCACTAACTG 1772
 |||||
 Db 2146 ACTCATGACCGCCCTCGTGCAGGCCCTGCGACGAGAGACGG 2192
 |||||

RESULT 11
 ID AAS06717 standard; cDNA; 2385 BP.
 AC AAS06717;
 XX
 DT 12-SEP-2001 (first entry)
 DE Polynucleotide sequence encoding human protein kinase #17.
 DE XX
 KW Human; protein kinase; PTK; STK; cancer; cardiovascular disease;
 KW metabolic disorder; immune related disease; neurological disorder;
 KW neurodegenerative disorder; inflammatory disorder; infectious disease;
 KW reproductive disorder; Gene therapy; ss.
 XX
 OS Homo sapiens.
 OS XX
 PN W0200138503-A2.
 XX
 PD 31-MAY-2001.
 XX
 PF 22-NOV-2000; 2000WO-US32085.
 XX
 PR 24-NOV-1999; 99US-0167482.
 XX
 XX (SUGEN-) SUGEN INC.
 PA
 XX Plowman GD, Whyte D, Manning G, Sudarsanam S, Martinez R;
 PI Flanagan P, Clary D;
 XX
 DR WPI; 2001-343950/36.
 DR P-PSDB; AAU03517.
 XX

Nucleic acids encoding human kinase polypeptides, useful for preventing
 diagnosing and/or treating e.g. cancer, immune, cardiovascular and
 neuronal-associated diseases, and microbial infections -
 Example 1; Figure 1; 433pp; English.

AAS06701-AAS06757 encode for novel human protein kinases #1-57. The
 novel protein kinases have been identified as members of the tyrosine
 or serine/threonine kinase (PK and STK) families. The polynucleotides
 encoding protein kinases and the polypeptides may be used in the
 prevention, diagnosis and treatment of diseases associated with
 inappropriate kinase expression. For example, they may be used to treat
 cancers (especially cancers of haematopoietic origin), cardiovascular
 disease (e.g. atherosclerosis), metabolic disorders (e.g. diabetes),
 immune related diseases (e.g. rheumatoid arthritis), neurological
 disorders (e.g. schizophrenia), neurodegenerative disorders (e.g.
 Parkinson's disease), inflammatory disorders (e.g. asthma), infectious
 disease (e.g. HIV) and reproductive disorders (e.g. infertility).
 Additionally, polynucleotides encoding protein kinases may be
 used for gene therapy and as DNA probes in diagnostic assays.
 The protein kinase polypeptides may be used as antigens in the production
 of antibodies against the protein kinases and in assays to identify
 modulators of protein kinase expression and activity.

Sequence 2385 BP; 494 A; 768 C; 731 G; 392 T; 0 other;
 Query Match 46.6%; Score 851.4; DB 22; Length 2385;
 Best Local Similarity 68.9%; Pred. No. 1.6e-176;
 Matches 1356; Conservative 0; Mismatches 416; Indels 195; Gaps 6;
 1 ATGAAGGTGGAGCGGAGATCGGACTCTGAGCTCATGAGCTCATGAGCAACCCACATGCTCTAAAG 60
 274 ATGAAGGTGGAGCGGAGATGCCATCTGAGCTCATGAGCAACCCACATGCTCTAAAG 333
 61 CTGACGACGCTTTATGAACCAAAAAATATTGTACTGCTGCTAGAACACGCTGTCAGGT 120

Db 334 CTCCACGAGCTCTACGAGACAGAATAATTGTACTGTTCTGGACACGCTCTCGGG 393
 |||||
 QY 121 GGTGAGCTCTTCGACTACCTGTTGAAGAAGGGAGAGCTGACGCTTAAGGAGGTCGGAAG 180
 |||||
 Db 394 GGTGAGCTATTTCGACTACCTGTTGAAGAAGGGAGAGCTGACGCTTAAGGAGGTCGGAAG 453
 |||||
 QY 181 TTCTTCGGCAGATCATCTCGCTGGACTTCTGCCACAGCCACATCCATATCCACAGG 240
 |||||
 Db 454 TTCTTCGGCAGATTTGTTCTGCTGGACTTCTGCCACAGCTACTCTCTCCACAGA 513
 |||||
 QY 241 GATCTGAACCTTGAACCTCTCTGCTGGACGAGAGAAACAACATCCGATCGCAGACTTT 300
 |||||
 Db 514 GACCTAAAGCCCGAGAACCTGCTTTTGGATGAGAAAAACAACATCCGATTCGAGACTTC 573
 |||||
 QY 301 GGATGGCGTCCCTGCGAGGTGGCGACAGCTTTGGAGACCAAGCTGTGGGTCCCCCAC 360
 |||||
 Db 574 GGATGGCGTCCCTGCGAGGTGGCGACAGCTTCTGGAGACCAAGCTGTGGGTCCCCCAT 633
 |||||
 QY 361 TAGCCTGCCCGGAGGTGATCCGGGGGAGAGATATGACGGCCGGAAGCGGAGCTGTGG 420
 |||||
 Db 634 TATGCGTGTCCAGAGGTGATTAAGGGGGAAAAATATATGATGGCCCGGCGAGACATGTGG 693
 |||||
 QY 421 AGCTGGCGGTCTATCTCTGTTGCTTGTGGTGGGGGCTCTGCTTTCGACGATGACAAC 480
 |||||
 Db 694 AGCTGTGGAGTCTATCTCTTGGCTGCTGCTGGGGGCTCTGCTTTCGACGATGACAAC 753
 |||||
 QY 481 TTGCGACAGCTGCTGGAGAAAGTGAAGGGGGGCTGTTTCAATGCGGCACTTTATCCCG 540
 |||||
 Db 754 CTCGCCAGCTGCTGGAGAAAGTGAAGGGGGGCTGTTTCAATGCGGCACTTTATCCCT 813
 |||||
 QY 541 CCGGACTGCCAGAGTCTGTACGGGGGATGATGAGGTGAGCGCCGACCGCCCTCAAG 600
 |||||
 Db 814 CAGATTGCCAGAGCTCTCTGAGGGGAATGATCGAAGTGGAGCCCGGAAAAAGGCTCAGT 873
 |||||
 QY 601 CTAGAGCACATTCAAGAAACACATATGTTATATAGGGGGCAAGAAATGAGCCCGAACCC 656
 |||||
 Db 874 CTGGAGCAAAATTCAGAAACATCTCTGTTAGTACCTAGGGGGGAAACACAGAGCCAGCCGTGC 933
 |||||
 QY 657 --AGAGAGCCCAATCTCTGCAAGGTGCGAGATCCGCTGCTGCCAGCTTGGAGGACATC 714
 |||||
 Db 934 CTGGAGCCAGCCCTTGGCGCGGTAGCCATCGGAGGCTGCTCCCAACCGAGAGCTG 993
 |||||
 QY 715 GACCCGAGCTGTGACAGCATGCTCACTCGCTGCTTCCGAGACCGCAACAGCTG 774
 |||||
 Db 994 GACCCGAGCTCTAGAGAGCATGGCATCACTGGCTGCTTCAAGGACCCGAGAGCTG 1053
 |||||
 QY 775 CTGACGAGCTGCTGTCCGAGGAGAGAACCCAGGAGAGATGATTTTACTTCTCTCCTCTG 834
 |||||
 Db 1054 CATCGGAGCTGCGCAGTGAGGAGGAGAACCAAGAAAGATGATATATATCTGCTTTTG 1113
 |||||
 QY 835 GACCGGAAGAAAGGTACCCGAGCCAGGAGGATGAGACCTGCCCTCCCGGAAACGAGATA 894
 |||||
 Db 1114 GATCGGAAGGAGCGGTATCCCACTGTGAGGACCCAGGACCTGCTCCCGGAAATGATGTT 1173
 |||||
 QY 895 GACCTCTCCCGAAGCGTGTGGACTCCCGATCTGAACCGGACGACGAGCGGCGGCA 954
 |||||
 Db 1174 GACCTCTCCCGAAGCGTGTGGATTCTCCATGCTGAGCGCTCAGGGGAAGCGGAGCA 1233
 |||||
 QY 955 GAACGCAAAATCCATGAGGTGCTCAGCGTGAACGGA-----CGCGGCTCCCCGGTG 1005
 |||||
 Db 1234 GAGCGGAAGTCCATGGAAGTCTCTGAGCATCACCGATCCGGGGGTGGTGGCTCCCCGTGA 1293
 |||||
 QY 1006 CTTGCGGGGGGGCCATTGAGATGGCCCGACGAGGCTCTGGTTCATCAGCGGT 1065
 |||||
 Db 1294 CCCACCGAGCGGCTTGGAGATGGCCCGACGACAGCCAGAGATCCCGTAGCGTCAAGTGA 1353
 |||||
 QY 1066 GCCTCTCAGGCTTTCCACCGCCACTCAGCAGCGCCCGGG----- 1108
 |||||
 Db 1354 GCCTCAGGCTCTGCTCTCCAGCCCTCTAAGCAGCCCAAGGAGTCCGGTCTTTTCTCTT 1413
 |||||
 QY 1109 ----- 1108


```
Db 274 ATGAAGGTGGAGCGGAGATCGCATCTCTGAAGTTCATCGAACCCACATGCTCTCAAG 333
Qy |||||
Db 61 CTGCACGAGCTTATGAAGAAACAAATATTTGTACCTGGTCTGCTAGAACACGTGTCAAGT 120
Db 334 CTCCACGAGCTCTACAGAAACAAAGAAATATTTGTACCTGGTCTGAGACACATCTCGGG 393
Qy |||||
Db 121 GGTGAGCTCTTCGACTACCTGTGTAAGAGGGGAGGCTGACGCCCTAAAGAGGCTCGGAAG 180
Db 394 GGTGAGCTATTGACTATCTGCTGTAAAGAGGGGAGACTGACGCCCAAGAGGCCCCGAAG 453
Qy |||||
Db 181 TTCTTCGGCAGATCATCTCTGCTGTGACTTTCTGCCACAGCACTTCCATATGCCACAGG 240
Db 454 TTCTTCGGCAGATCTGCTGCTGTGACTTTCTGCCACAGCTACTCCATCTGCCACAGA 513
Qy |||||
Db 241 GATCTGAACCTTGAACCTCTCTGCTGAGCAGAGAGAACAAATCGGCATCGCAGACTTT 300
Db 514 GACCTAAAGCCCCGAGAACCTGTCTTTGGATGAGAAACAAATCGCATTTGACAGCTTC 573
Qy |||||
Db 301 GGCATGGCGTCTCTGAGGTTCGCCACAGCCTGTTGGAGACAGCTGTGGGTCCCCCCAC 360
Db 574 GGCATGGCGTCTCTGAGGTTCGCCACAGCCTCTCTGGAGACAGCTGTGGGTCCCCCCAT 633
Qy |||||
Db 361 ATAGCCTGCCCCGAGGTGATCGGGGGGAGAAATATGACGGCCGGAAGCGGACGTGTGG 420
Db 634 TATGCGTGTCCAGAGGTGATTAAGGGGGGAAATATGATGGCCGCGGCGAGACATGTGG 693
Qy |||||
Db 421 AGCTGGGGGTCTATCTGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 480
Db 694 AGCTGTGGAGTCTATCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 753
Qy |||||
Db 481 TTGGCAGACGCTCTGAGAGGTGAAGCGGGCGGTGTTCCATATGCCGACCTTTATCCCG 540
Db 754 CTCCGCCAGCTGCTGGAGAGGTGAACCGGGCGCTCTTCCATGCTCCCATCTTCACTTCT 813
Qy |||||
Db 541 CCGCAGCTGCGAGAGTCTGCTACGGGGCATGATCGAGGTGGACCGCGCACCGCGCTCACG 600
Db 814 CCAGATTGCCAGAGCTCTCTGAGGGGAATGATCGAAGTGGAGCCCGGAAAGGCTCAGT 873
Qy |||||
Db 601 CTAGAGCACTTCAGAAACATATGTTATATAGGGGGCAAGATGAGCCCGAAC----- 656
Db 874 CTGGAGCAAAATTCAGAAACATCTTTGGTACCTTAGCGGGGAAACACAGGCCAGACCCGTGC 933
Qy |||||
Db 657 --AGAGCAGCCCATCTCTCGAAGGTGCAGATCCGCTGCTGCCAGCTTGAGGACATC 714
Db 934 CTGGAGCCAGCCCTTGGCGCGGTAGCCATGCGAGCCTGCCATCAAAGAGAGCTG 993
Qy |||||
Db 715 GACCCCGAGCTGCTGGAAGCATGCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 774
Db 994 GACCCCGAGCTCTAGAGAGCATGCGCATCACTGGGCTGCTTCAGGGACCGCGAGAGGCTG 1053
Qy |||||
Db 775 CTGAGGACCTGCTGCTGAGAGAGAGAACAGAGAGAGATGATTTACTTCTCTCTG 834
Db 1054 CATCCGAGCTGCGCAGTGAGAGAGAGAACCAAGAAAGATGATATATTTATCTGCTTTG 1113
Qy |||||
Db 835 GACCCGAAAGAGGTACCCGAGCAGAGAGATGAGGACCTGCCCGCCCGGAAACAGAGATA 894
Db 1114 GATCGGAGAGCGGTATCCAGCTGTGAGGACACAGGACTGCTGCCCTCCCGGATGATGTT 1173
Qy |||||
Db 895 GACCTTCCCCGAGAGCTGTGACTCCCCGATGCTGGAACCGGACAGCGCAAGCGCGGCA 954
Db 1174 GACCCCGCCCGGAGCGTGTGATTTCTCCCATGCTGAGCGGTCAAGGAGCGCGGACCA 1233
Qy |||||
Db 955 GAAACCAATTCATGGAGGTGCTACGCTGAGCGA-----CGGCGGTCCCCGCTG 1005
Db 1234 GAGCGGAAGTCCATGGAAGTCTGAGCATACACCGATGCGCGGGGTGCTGCTCCCTGTA 1293
Qy |||||
Db 1006 CTTGCGCGCGCGGCTTATGAGATGCGCCAGCAGCGCTCAGAGTCTCGGTTCATACAGCGGT 1065
Db 1294 CCAACCCGAGCGGCTTGGAGATGGCCAGCAAGCAGAGATCCGTAGCGTCAAGTGA 1353
Qy |||||
Db 1066 GCCTCTCTAGGCTTTTCCACAGCCCACTCAGACGCCCGCGG----- 1108
Qy |||||
```

RESULT 13

ABV74558

ID ABV74558 standard; cDNA; 2289 BP.

XX AC ABV74558;

XX DT 20-JAN-2003 (first entry)

XX DE Human kinase #2 coding sequence.

XX XX Human; kinase; chromosome 19; gene; ss.

XX OS Homo sapiens.

XX

```
Db 1354 GCCTCAAGGGTCTGTCTCTCCAGCCCTTAAAGCAGCCCAAGAGTCCGGTCTTTTCTCTTT 1413
Qy |||||
Db 1109 ----- 1108
Db 1414 TCACCGGAGCGGGGCTGGAGATGAGGCTCGAGCGGGGCTCCCGACTTCCAAACG 1473
Qy |||||
Db 1109 -----TGACCCCTCACCCC 1122
Db 1474 CAGAGCTGCTTCTCGGGGCCCCAGGGTGGGGGCGCGGGGAGACGCCCGCCGCCCCC 1533
Qy |||||
Db 1123 TCACAAAGGGCAGTCCCTTCCCC-----ACC 1149
Db 1534 AGTCCGCTCCACACCCCTGCGCGCCCCCAGGCTCCCGCGCTCTCTCTGCGGGAGC 1593
Qy |||||
Db 1150 CCAAGGGGACACTGTCTCACACGCCCAAGAGAGAGCCGGCTGGCAGCCCAACCCACG 1209
Db 1594 CCCTTGCACTCGCTCTGCAACGCCCGGGGCCAGTCCCAACCGGAGCCCGGGGACACA 1653
Qy |||||
Db 1210 CCCCCTCCAGCCCC-----AGCGTCGAGGGGTGCCCTGGAGGGCGCGGCTCAACTCC 1263
Db 1654 CCACCCCGACCCCGGGGTGGGCTCGGGGAGCCGCTGGAGAGTCTGCTCAACTCC 1713
Qy |||||
Db 1264 ATCAAGAACAGCTTTCTGGCTCACCCCGCTTCCACCGCCGGAACCTGCAAGTTCCGACG 1323
Db 1714 ATCCGCAACAGTTCCTGGGCTCCCTCGCTTTCACCGGCGCAAGATGACAGTCCCTACC 1773
Qy |||||
Db 1324 CCGAGAGAGATGTCACACCTGACACAGAGTCTGCCAGAGCTGGGAGAGAAAGTCTCTGG 1383
Db 1774 GCTGAGGAGATGTCAGCTTGAAGCCAGAGTCTCTCCCGAGAGTGGCAAAACGCTCTCTGG 1833
Qy |||||
Db 1384 TTTGGGAACCTTCATCAGCTGGAGAGGAGAGCAGATCTTCGTGGTTCATCAAAAGACAAA 1443
Db 1834 TTCGGGAACCTTCATCTCTTGGACAAAGAGAACAAATATTCCTCGTGTAAAGACAAA 1893
Qy |||||
Db 1444 CCTCTGAGCTCCATCAAGGTGACATCGTCACGCCCTCTCTGTCGATTCACAGTCTCAGC 1503
Db 1894 CCTCTCAGCAGCATCAAAAGCAGACATCGTCCATGCTTCTGTGATCCCGAGCTGAGT 1953
Qy |||||
Db 1504 CACAGCTCATCTCCAAACGAGCTTCGGGGCGGAGTACAAGGCCACGGGGGGGCCAGCC 1563
Db 1954 CACAGTGTGTGTACAGACCAAGCTTCAGGGCGGAGTACAAGGCCAGTGGCGGCCCTCC 2013
Qy |||||
Db 1564 GTGTTCAGAAAGCGGTCAAGTTCACAGTGTGATATCACCTACACGGAGGT----- 1614
Db 2014 GTCTTCCAAAGCCCGTCCGCTTCAGGTGACATCAGCTCCTCTGAGGGTCCAGAGCCC 2073
Qy |||||
Db 1615 -----GGGAGGCGCAGAGAGAAAGGCAATCTACTCCGTCACCTTACCCCTGCTC 1665
Db 2074 TCCCGCGAGCGGACGCGAGGTGGTGGCATCTACTCCGTCACCTTCACTCTCATC 2133
Qy |||||
Db 1666 TCAGGCCCGAGCGCTGCTTCAGAGGGTGGTGGAGAGCCATCCAGGCCAGCTGCTGAGC 1725
Db 2134 TCGGGTCCAGCGCGTCCGTTCAAGCGAGTGGTGGAGACCATCCAGGACACAGCTCTGAGC 2193
Qy |||||
Db 1726 ACACAGACCCCGCTGGGCGCCAGCACTTGTTCAGACACCACTAACTG 1772
Db 2194 ACTCATGACAGCCCTCGGTGACGGCCCTGGCAGCAGAGAACGG 2240
```

PH	Key	Location/Qualifiers
FT	CDS	1..2289
FT		/*tag= a
FT		/product= "Human kinase #2"
XX	W0200281670-A1.	
XX	PN	
XX	PD	17-OCT-2002.
XX	XX	
XX	PF	04-APR-2002; 2002MO-US10786.
XX	PR	06-APR-2001; 2001US-282036P.
XX	XX	(LEXI-) LEXICON GENETICS INC.
XX	PA	
XX	PI	Turner CA, Mathur B, Friddle CJ;
XX	DR	WPI; 2003-058538/05.
XX	DR	P-PSDB; ABB98744.
XX	XX	
PT	PT	New human kinase proteins useful for diagnosis, drug screening,
PT	PT	clinical trial monitoring, treatment of disorders and diseases, and
PT	PT	cosmetic and nutritional applications
XX	XX	
XX	PS	Claim 5; Page 41-42; 47pp; English.
XX	XX	
CC	CC	The present sequence is the coding sequence for a novel human kinase. The
CC	CC	genomic locus encoding the kinase is thought to be on human chromosome
CC	CC	19. The kinase and its coding sequence are useful for diagnosis, drug
CC	CC	screening, clinical trial monitoring, treatment of disorders and
CC	CC	diseases, and cosmetic and nutritional applications.
XX	XX	
SQ	SQ	Sequence 2289 BP; 457 A; 759 C; 666 G; 407 T; 0 other;
Query Match 42.5%; Score 777.2; DB 25; Length 2289;		
Best Local Similarity 67.9%; Pred. No. 2.6e-160;		
Matches 1277; Conservative 0; Mismatches 408; Indels 195; Gaps 6;		
QY	88	TATTTCTACTGCTGTAGAACACGCTGTCAAGTGTGTGAGCTTTCGACTACCTGGTGAAG 147
DB	265	TTTAGGTACCTGTTCTGGAGCAGCTTCGGGGGTGAGCTATTCGACTACCTGGTAAAG 324
QY	148	AAGGGAGGCTGACCGCTAAGGAGGCTCGGAAGTTCTTCGGCAGATCATCTCTGGCGTG 207
DB	325	AAGGGAGACTGACGCCCAAGAGGAGCCGAAAGTTCTTCGGCAGATGTGCTCGCGTG 384
QY	208	GACTTTCGCCACAGCCACTCCATATGCCACAGGATCTGAAACCTGAAACCTCTGCTG 267
DB	385	GACTTTCGCCACAGCTACTCCATCTGCCACAGAGACCTAAAGCCCGAGAACCTGCTTTG 444
QY	268	GACGAGAGAACACATCCGATCGCAGACTTTGGCATGGCGTCCCTGCAGGTTGGGAC 327
DB	445	GATGAGAAAAACAATCCGATTCGACACTTCGGCATGGCGTCCCTGCGAGGTGGGGAC 504
QY	328	AGCTGTGTGAGAGCAAGCTGTGGTCCCCCACTATCGCTGCCCGAGGATGATCCGGGG 387
DB	505	AGCTTCTGAGAGCAAGCTGTGGTCCCCCACTATCGCTGCCCGAGGATGATTAAGGG 564
QY	388	GAGAAGTATGACGCCCGGAGAGGCGGAGCTGTGAGCTGCGCGTCACTCTGCTTGCCTTG 447
DB	565	GAATAATATGATGGCCCGCGGAGACATGTGGAGCTGTGGAGTCACTCTTTCGCCCTG 624
QY	448	CTGTTGGGGCTGCGCTTCGAGTGAACATCTGGCAGAGCTGTGGAGAGGTTGAAG 507
DB	625	CTGTTGGGGCTGCGCTTTCGAGTGAACATCTGGCAGAGCTGTGGAGAGGTTGAAG 684
QY	508	CGGGCGGTGTTCACATCGCGCACTTTATCCCGCCGACTGCCAGAGTCTGCTACGGGGC 567
DB	685	CGGGCGGTGTTCACATCGCGCACTTTATCCCGCCGACTGCCAGAGTCTGCTACGGGG 744
QY	568	ATGATCAGGTGAGCGCGCACTTCAGCTCAGCTAGACACATTCAGAAACATATGG 627
DB	745	ATGATCAGGTGAGCGCGCACTTCAGCTCAGCTAGACACATTCAGAAACATATGG 804
QY	628	TATATAGGGGCAAGATGAGCCCGAAC-----AGAGAGCCCATTCCTCGCAAGGTG 681
DB	805	TACCTTAGCGGGAACACAGAGCCAGACCGCTGCTGGAGCCAGCCCTGGCGCGGGTA 864
QY	682	CAGATCCGCTCGCTGCCAGCCTGAGAGACATCGACCCGAGCTGCTCGACAGCATGCAC 741
DB	865	GCCATGCGGAGCTGCCATCAACGAGAGCTGGACCCGACGCTCTAGAGAGCATGGCA 924
QY	742	TCATGGGCTGCTCCGAGACCGCAACAAAGTGTGTGAGGACCTGCTGTCCGAGAGGAG 801
DB	925	TCATGGGCTGCTTCAGGAGCCGCGAGAGGCTGCTCGGAGCTCGCAGTGGAGAGGAG 984
QY	802	RACCAGGAGAGATCATTTACTTCTCTCTCTGGACCGGAAGAGGTACCCGAGCCAG 861
DB	985	AACCAAGAAAGATGATATATATCTGCTTTTGGATCGGAAGAGCGGTATCCCCAGTGT 1044
QY	862	GAGGATGAGGACCTGCCCGCCCGGAAACGAGATAGACCTCCCGGAAAGCTGTGACTCC 921
DB	1045	GAGGACGAGGACCTGCCCGGAAAGTGTGAGTGTGACCCCGGAGAGGCTGTGATTTCT 1104
QY	922	CCGATGCTGAACCGGACCGCAACGCGGCGGAGCAACGCAATTCATGAGGTGTCTCAGC 981
DB	1105	CCCATGCTGAGCGGTACGGAAGCGGCGACAGAGCGGAAGTCCATGGAAGTCTCTGAGC 1164
QY	982	GTGAC-----CGACGGGGCTCCCGGTGCTCGCGGGGGCCATTGAGATGGCC 1032
DB	1165	ATCACCGATGCGGGGGTGGTGGCTTCCCTGTACCCACCGACGGGGCTTTGAGATGGCC 1224
QY	1033	CAGACGCGCAGAGGTCTCGGTCCATCAGCGGTGCTCTCTCAGGCTTTTCCACACAGCCCA 1092
DB	1225	CAGCACAGCCAGAGATCCGTTAGCTCAGTGAGGCTTCACGGGTCTGTCTCCAGCCCT 1284
QY	1093	CTCAGCAGCCCGC----- 1106
DB	1285	CTAAGCAGCCCAAGAGTCCGCTTTCTTTTCCCGGAGCGGGGCTGGAGATGAG 1344
QY	1107	----- 1106
DB	1345	GCTCAGGGGGGGCTCCCGACTTCGAAACGAGAGAGCTGCTCTTCGCGGGCCCCAGG 1404
QY	1107	-----GGTGACCCCTCACCCCTCACCAAGGGGAGTCCCTCCCTCCCT----- 1146
DB	1405	GGTGGGGCGCGGGGAGCAGCCCCCGCCCCAGTCCCGCTCCACACCTCTGCGCGGC 1464
QY	1147	-----ACCCCAAGGGGACACTGTCTCACAGCCCA 1176
DB	1465	CCCCCAGGCTCCCGGCTCTCTGCGGGGACCCCTTGCCTCGCTCTGCACAGCCCT 1524
QY	1177	NAGGAGAGCCCGCTGGCAGCGCCCAACCCCGCCCGTCCAGCCCT-----AGCGTC 1230
DB	1525	CGGGCCAGTCCACCGGAGCCCCGGGGAACACACACCCCGCCCGCGGGTGGCGTC 1584
QY	1231	GGAGGGGTGCCCTGGAGGGCGGGCTCAACTCCATCAAGAAACAGCTTTCTGGGCTCACCC 1290
DB	1585	GGGGAGCGCGCTGGAGGAGTCTCTCACTCCATCCGCAACAGCTTCTGCGGCTCCCT 1644
QY	1291	CGCTTCCACCGCGGAAACTGCAAGTTCCGAGCGCGGAGGAGATGTCCAACTGACACCA 1350
DB	1645	CGCTTTTACCGGCGCAAGATGCAAGTCCCTACCGCTGAGGAGATGTCCAGCTTGACGCCA 1704
QY	1351	GAGTGTCTCCAGAGCTGGCGAAGAGTCTCTGTTTGGAACTTCATCAGCTCGGAGAG 1410
DB	1705	GAGTCTCTCCCGAGCTGGCAAAACGCTCTCTGTTTGGGAACTTCATCTCTTGGACAA 1764
QY	1411	GAGGAGCAGATCTTCGTGGTTCATCAAGACAAACCTCTGAGCTTCCATCAAGGTGACATC 1470
DB	1765	GAAGAACAAATATTCCTGCTGCTAAAGGACAAACCTCTCAGCAGCATCAAGCAGACATC 1824
QY	1471	GTGCAGCCCTTCTGTGATTCCTCAGTCTCAGCCAGAGCTCATCTCCCAAGAGCTTC 1530
DB	1825	GTCCATGCTTTCTGTGATTCCTCCAGCTCAGTGTGCTGTCAAGAGCCAGCTTC 1884

QY 1531 CGGCGCAGTACAAAGCCAGCGGGGGCCAGCGGTGTTCAGAAAGCGGTCAAGTTCCAG 1590
 Db 1885 AGGCGCCAGTACAAAGCCAGTGGCGGCCCTCGTCTTCCAAAAGCCGCTCCGCTTCCAG 1944
 QY 1591 GTTGATATCACTACAGGAGGT-----GGGAGGCGCAGAAGGAG 1632
 Db 1945 GTGGACATACGCTCTCTGAGGTTCAGAGCCCTCCCGGAGCGGAGCGGAGGT 2004
 QY 1633 AAGCGCATCTACTCCGTCACCTTCAACCTCTCTCAGGCGCCAGCGGTCTTCAAGAGG 1692
 Db 2005 GGTGGCATCTACTCCGTCACCTTCACTCTCTCTGCTCCAGCGGTCTTCAAGCA 2064
 QY 1693 GTGGTGAGACCATCAGGCGCCAGCTGTGTGAGCACACAGCCGCTCGGCGCCAGCAC 1752
 Db 2065 GTGGTGAGACCATCAGGCGCCAGCTGTGTGAGCACACAGCTCATGACCTCATGACCGCTCGGTGAGGCC 2124
 QY 1753 TTGTGACAGACCACTAATG 1772
 Db 2125 CTGGCAGAGAGAAACGG 2144

RESULT 14

ABX70994

ID ABX70994 standard; cDNA; 1797 BP.

XX AC ABX70994;

XX XX

DT 05-MAR-2003 (first entry)

XX DE Novel human cDNA sequence #219.

XX KW

KW Human; gene; ss; nervous system disorder; peripheral neuropathy;
 KW Huntington's disease; amyotrophic lateral sclerosis; haemophilia;
 KW neurodegenerative disease; Parkinson's disease; Alzheimer's disease;
 KW autoimmune disease; systemic lupus erythematosus; rheumatoid arthritis;
 KW insulin-dependent diabetes mellitus; anaemia; thrombocytopaenia; wound;
 KW ulcer; burn; bone disorder; osteoporosis; osteoarthritis; stroke;
 KW fibrosis; reperfusion injury; infection; allergic rhinitis; asthma;
 KW coagulation disorder; cancer; tumour; inflammatory disease;
 KW septic shock; Crohn's disease; anaphylaxis; proliferation; chemotactic;
 KW differentiation; stem cell growth factor; haematopoiesis; chemokinetic;
 KW haemostatic; antiinflammatory; expressed sequence tag; EST.

XX OS Homo sapiens.

XX XX

XX PN WO200281731-A2.

XX XX

XX PD 17-OCT-2002.

XX XX

XX PF 29-JAN-2002; 2002WO-US01222.

XX XX

XX PR 30-JAN-2001; 2001US-0774528.

XX XX

XX PA (HYSE-) HYSEQ INC.

XX PA (GOOD/) GOODRICH R. W.

XX XX

PI Tang TY, Liu C, Zhou P, Asundi V, Zhang J, Zhao QA, Ren F;

PI Xue AJ, Yang Y, Wehrman T, Wang J, Wang D, Drmanac RT;

XX XX

XX DR WPI; 2003-058563/05.

XX XX

XX Novel polypeptide useful for treating neurodegenerative diseases,
 XX myeloid or lymphoid cell disorders, bone disorders, mechanical and
 XX traumatic disorders, coagulation disorders, and inflammatory diseases

XX PT

XX PS Claim 1; Page -; 612pp; English.

XX XX

CC This invention relates to the cDNA sequences encoding an isolated
 CC novel human polypeptide. The protein encoded by the nucleic acid of
 CC the invention is useful for treating central and peripheral nervous
 CC system diseases (e.g. peripheral neuropathy, Huntington's disease,
 CC amyotrophic lateral sclerosis); neurodegenerative diseases (e.g.

CC Parkinson's disease, Alzheimer's disease); autoimmune disease (e.g.
 CC systemic lupus erythematosus, rheumatoid arthritis, insulin-dependent
 CC diabetes mellitus); myeloid or lymphoid cell disorders (e.g. anemia
 CC and thrombocytopaenia); wounds, ulcers, burns; bone disorders (e.g.
 CC osteoporosis, osteoarthritis); mechanical and traumatic disorders (e.g.
 CC stroke, head trauma); lung or liver fibrosis; reperfusion injury in
 CC various tissues; bacterial, viral or fungal infections; allergic
 CC conditions such as allergic rhinitis, asthma; coagulation disorders
 CC (e.g. haemophilia); cancer and tumours; and inflammatory diseases (e.g.
 CC septic shock, Crohn's disease, anaphylaxis). The protein may be used to
 CC inhibit the growth, infection or function of infectious agents such as
 CC bacteria, fungi, viruses, or to effect bodily characteristics,
 CC biohythms or circadian cycles of rhythms. The protein may also
 CC have proliferative/differentiation, stem cell growth factor,
 CC haematopoiesis regulation, immune stimulating or suppressing,
 CC chemotactic/chemokinetic, haemostatic and thrombolytic, receptor/ligand,
 CC and antiinflammatory activities. The cDNA sequences of the invention are
 CC useful for expressing recombinant protein for analysis. The present
 CC sequence represents a novel human cDNA sequence of the invention.
 CC this sequence is an expressed sequence tag (EST) and was identified
 CC using subtractive hybridisation.

XX SQ Sequence 1797 BP; 375 A; 549 C; 520 G; 353 T; 0 other;
 Query Match 32.2%; Score 588.8; DB 25; Length 1797;
 Best Local Similarity 77.1%; Pred. No. 4e-119;
 Matches 748; Conservative 0; Mismatches 207; Indels 15; Gaps 2;

QY 93 GTACCTGTGTAGAACACGTGTAGGTGGTGGTCTTTCGACTACCTGGTGAAGAGG 152
 Db 89 GTACCTGTGTCTGGAGCACGTCTCGGGGGTGCAGTATTTCGACTACCTGGTGAAGAGG 148
 QY 153 GAGGCTGACGCTTAGAGGAGGTTCGAAGTCTTTCGGCAGATCACTCTCGGCTGGATT 212
 Db 149 GAGACTGAGCCCAAGAGAGGCCGAAAGTCTTTCGCCAGAGTTGTTCTGCGCTGGATT 208
 QY 213 CTGCCACAGCCACTCATATGTCACAGGGATCTGAAACCTGAAACCTCTCTCGAGCA 272
 Db 209 CTGCCACAGCTACTTCATCTGCCACAGAGACTTAAAGCCCGAGACCTGCTTTGGATGA 268
 QY 273 GAAGAACAAATCCGCATCGCAGACTTTGGGATGCGCTCCCTGCGAGTTGGCGACAGCCT 332
 Db 269 GAAAAACAACATCCGCATTGCGACTTTCGAGACTTCGCGTCCCTGCGAGTTGGCGACAGCCT 328
 QY 333 GTTGGAGACCAAGCTGTGGTCCCGCCACTAGCCCTGCCCGAGGTGATCCCGGGGAGAA 392
 Db 329 CCTGGAGACCAAGCTGTGGTCCCGCCACTATGCTGTCCAGAGGTGATTAAGGGGAGAAA 388
 QY 393 GTATGACGGCCGGAAGCGGACGTGTGGAGCTGCGCGCTCATCTCTGTTCTGCTTCTGCT 452
 Db 389 ATATGATGGCCCGCGGCGAGACATGTGGAGCTGTGAGTCTATCTCTTCCGCTTCTGCT 448
 QY 453 GGGGCTCTGCCCTTCGACGATGACAACTTGGGACAGCTGTCTGGAGAGAGGTGAAGCGGG 512
 Db 449 GGGGGCTCTGCCCTTTGATGACGACAACTTCGCCAGCTGTCTGGAGAGAGGTGAAGCGGG 508
 QY 513 GCTGTTCCACATGCGCAGCTTTATCCCGCCGAGCTGCGAGAGTCTGCTACGGGGCATGAT 572
 Db 509 GCTTTCACATGCCCCCACTTCATTCCTCCAGATTGCCAGAGCTCTCTGAGGGGAATGAT 568
 QY 573 CGAGTGGAGCGCCGACCGCCCTCACGCTAGAGCACATTCAGAAACACATATGTTATAT 632
 Db 569 CGAGTGGAGCGCCGAAAAAGGCTCAGTCTGGAGCAATTCAGAAACATCTCTTGGTACT 628
 QY 633 AGGGGGCAAGAAATGAGCCCGAACCA-----GAGCAGGCCATTCCTCGCAAGTGCAGAT 686
 Db 629 AGGGGGCAACACAGGACCGACCGCCCTGCTGTGAGCCAGCCCTCGCGCGGGTGAACCAT 688
 QY 687 CGGCTCGCTGCCCGACCTGGAGGACATCGACCCGAGCTGTGGACAGCATCTCACTACT 746
 Db 689 GGGGAGCTTGCATCATCAACGGAGAGCTGGAGCCCGACGCTCTTAGAGAGATGGCATCACT 748
 QY 747 GGGCTGTCTTCCGAGACCGCAACAGCTGTCTGAGGACCTGTGTCTCCGAGGAGGAGAACCA 806

Db 749 GGGCTGCTTCAGGACCGGAGAGGCTGCATCGAGCTGCGAGTGCGAGTGAGGAGAGAACCA 808
Qy 807 GGAGAGATGATTTACTTCTCTCTGACCGGAAAGAGATACCGGAGCCAGGAGGA 866
Db 809 AGAAAGATGATATATATCTGCTTTTGGATCGGAGGAGCGTATCCAGCTGTGAGGA 868
Qy 867 TGAGGACCTGCCCCCGGAGAGATAGACCTTCCCGGAGGCTGTGACTTCCCGAT 926
Db 869 CCAGGACCTGCTCCCGGAGATGATGTTGACCCCCCGGAGGCTGTGATTTCTCCAT 928
Qy 927 GCTGAACCGGACGCGAGCGGCGGACGAGACGAAATCCATGAGGTGCTCAGCGTGAC 986
Db 929 GCTGAGCGCTCAGCGGAGCGGCGGACGAGGAGTCCATGGAAGTCTCTGAGCATCAC 988
Qy 987 -----GGAGCGCGCTCTCCCGGTGCTGCGCGCGGCGCATTTGAGATGGCCCCAGCA 1037
Db 989 CGATGCGGGGGTGGTGGCTCCCTGTATCCACCGACCGGCGCTTGGAGATGGCCCCAGCA 1048
Qy 1038 CGGCGAGAGG 1047
Db 1049 CAGCCAGAGG 1058

RESULT 15

ABL10489
ID ABL10489 standard; cDNA; 2720 BP.
AC ABL10489;
XX
XX
DT 26-MAR-2002 (first entry)
XX
DE Drosophila melanogaster expressed polynucleotide SEQ ID NO 25949.
XX
XX
KW Drosophila; developmental biology; cell signalling; insecticide;
KW pharmaceutical; gene; ss.
XX
OS Drosophila melanogaster.
XX
FN WO200171042-A2.
XX
PD 27-SEP-2001.
XX
PF 23-MAR-2001; 2001WO-US09231.
XX
XX
PR 23-MAR-2000; 2000US-191637P.
PR 11-JUL-2000; 2000US-0614150.
XX
XX
PA (PEKE) PE CORP NY.
XX
XX
PI Venter JC, Adams M, Li PWD, Myers EW;
XX
XX
DR WPI; 2001-656860/75.
DR P-PSDB; ABB66386.
XX
XX
PT New isolated nucleic acid detection reagent for detecting 1000 or more
genes from Drosophila and for elucidating cell signalling and cell-cell
interactions -
XX
XX
PS Claim 1; SEQ ID NO 25949; 21pp + Sequence Listing; English.
XX
XX
CC The invention relates to an isolated nucleic acid detection reagent
capable of detecting 1000 or more genes from Drosophila. The invention is
useful in developmental biology and in elucidating cell signalling and
cell-cell interactions in higher eukaryotes for the development of
insecticides, therapeutics and pharmaceutical drugs. The invention
discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA
sequences (ABL01840-ABL16175) and the encoded proteins
(ABB57737-ABB72072).
CC
CC The sequence data for this patent did not form part of the printed
specification, but was obtained in electronic format directly from WIPO
at ftp.wipo.int/pub/published_pct_sequences.
XX
XX

SQ Sequence 2720 BP; 661 A; 715 C; 759 G; 585 T; 0 other;
Query Match 24.4%; Score 446.2; DB 23; Length 2720;
Best Local Similarity 70.2%; Pred. No. 6.8e-88;
Matches 614; Conservative 0; Mismatches 258; Indels 3; Gaps 1;
Qy 1 ATGAAGTGGAGCGGAGATCGGATCTCGAGCTCTGAAGCTCATTGAGCACCCACCTCTTAAG 60
Db 178 ATGAAGTGGAGCGGAGATCGGATCTCGAGCTCTGAAGCTCATTGAGCACCCACCTCTTGC 237
Qy 61 CTGACAGAGCTTTATGAAAACAAAATAATTTGTACTCTGCTGTAGAACACAGTGTGAGT 120
Db 238 CTGAGGATGTGTACGAGAACAGAGTATTTGTATTTGATATTGAGGATGTATCCGCG 297
Qy 121 GGTGAGCTTTTCGACTACTCTGTGTGAAGAGGAGGAGCTGACGCTTAAGAGGCTCGAAG 180
Db 298 GGAGAGCTCTTCGATTTACCTGTGTAAGAGGCTCGATTGACGCGAAGAGGCGCGCAAG 357
Qy 181 TTCTTCCGCGAGATCATCTCTGGCTGGACTTCTGCCACAGGCACTCCATATGCCACAGG 240
Db 358 TTCTTCCAGCAATCATCTCTGGCTGGACTTCTGCCACTCGCATTCGATTTGCCATCGC 417
Qy 241 GATCTGAAAACCTGAAAACCTCTCTGTGACGAGAGAACAAACATCCGATCGCAGACTTT 300
Db 418 GACTTGAAGCGGAGATCTCTGTCTGGAGAGAGAGATAACATTAAGATAGCGGACTTT 477
Qy 301 GGATGCGCTCTCTGAGGTGGCGACAGCTTGTGGAGACAGCTGTGGGTCTCCGCCAC 360
Db 478 GGGATGGCTTCCCTGCGACGAGCTGGCAGCATGTTGGAGACCTCTCTGCGGAGCCACAC 537
Qy 361 TAGCGCTGCCGAGGTGATCCGGGGGAGAGATGAGCGCGGAGGCGGAGCTGTGG 420
Db 538 TAGCGCTGTCTAGAGGTCTACGCGGCGAGAGTACGATGGCGCGAAGCGGATGTCTGG 597
Qy 421 AGTGGCGGCTCATCTCTGCTGTGCTGGTGGGGCTCTGCGCTTCGACATGACAAAC 480
Db 598 TCCTGTGGGGTCTATCTCTATGCGCTCTCTGTGGTGGCTGCGCTTCGACGAGCAAC 657
Qy 481 TTGCGACAGCTGTGTGAGAAAGTGAAGCGGGGCTGTTCACATGCGGCACTTTATCCCG 540
Db 658 TTGCGCGAGCTGTGTGAGAAAGTCAAGCGGGGCTCTTTTCAATACCGCACTTTGTGCG 717
Qy 541 CCGGACTGCCAGAGTCTGCTAGCGGGCATGATCGAGGTGAGCGCGCATCGCGCTCACG 600
Db 718 CCGGACTGCCAGAGTCTGCTGCGGCGATGATTGAGGTCAATCCGAGACCGCGGCTCACG 777
Qy 601 CTAGAGCACATTCAGAAACACATATGTTATATA ---GGGGGCAAGAAATGAGCCCGAACCA 657
Db 778 CTGGCTGAAATCAACCGGCATCCGTGGGTCAAGCTGGCGCAAGGGGAGCTGGAGCTG 837
Qy 658 GAGCAGCCCATTCCTCGAAGGTGCGAGATCCGCTGCTGCCGAGCTGGAGGACATCGAC 717
Db 838 GAGCTGCCAATGATGGAGGTGGTGCAGACACACGTTTATTTCCACACCCCGCGGTGGAT 897
Qy 718 CCGGAGCTGTGTGAGACAGTCACTCACTGGGTGCTTCCGAGACCGCAACAGCTGTG 777
Db 898 CCGGATGTGTGAAACCGGATTTGCTCGCTGGCTGTTTCAAGGAGAGGAGAACTCATC 957
Qy 778 CAGGACCTGTGTCCGAGGAGGAGAACACAGGAGAGATGATTTACTTCTCTCTCTGGAC 837
Db 958 CAGGAACTGTCTAGTTCAAGTCAATACGAGAGGTTATATATTTCTCTGCTCGAG 1017
Qy 838 CGGAAAGAGGTACCCGAGCCAGGAGGATGAGGA 872
Db 1018 CGCAACGAGAGACGACCTCGCTGGAGGATGATGA 1052

Search completed: November 27, 2003, 05:15:03

Job time : 472.512 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model
Run on: November 27, 2003, 02:00:45 ; Search time 6538.89 Seconds
(without alignments)
11430.368 Million cell updates/sec

Title: US-10-054-579-3
Perfect score: 1827
Sequence: 1 atgaaggtggagcggagat.....gaattatcccgaaagttaa 1827

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 288711 seqs, 20454813386 residues
Total number of hits satisfying chosen parameters: 5777422

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : GenEmbl:

- 1: gb_ba.*
- 2: gb_hgt.*
- 3: gb_in.*
- 4: gb_om.*
- 5: gb_ov.*
- 6: gb_pat.*
- 7: gb_ph.*
- 8: gb_pl.*
- 9: gb_pr.*
- 10: gb_ro.*
- 11: gb_ste.*
- 12: gb_sy.*
- 13: gb_un.*
- 14: gb_vi.*
- 15: em_ba.*
- 16: em_fun.*
- 17: em_hum.*
- 18: em_in.*
- 19: em_mui.*
- 20: em_om.*
- 21: em_or.*
- 22: em_ov.*
- 23: em_pat.*
- 24: em_ph.*
- 25: em_pl.*
- 26: em_ro.*
- 27: em_ste.*
- 28: em_un.*
- 29: em_vi.*
- 30: em_htg_hum.*
- 31: em_htg_inv.*
- 32: em_htg_other.*
- 33: em_htg_mus.*
- 34: em_htg_pln.*
- 35: em_htg_rod.*
- 36: em_htg_mam.*
- 37: em_htg_vrt.*
- 38: em_sy.*
- 39: em_htgo_hum.*
- 40: em_htgo_mus.*
- 41: em_htgo_other.*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB	ID	Description
1	1827	100.0	1827	6	AX661193	AX661193 Sequence
2	1827	100.0	2007	6	AX661191	AX661191 Sequence
3	1825.4	99.9	2908	6	AR232170	AR232170 Sequence
4	1825.4	99.9	3516	9	AY166857	AY166857 Homo sapi
5	1813	99.2	3364	6	AR232171	AR232171 Sequence
6	1759.6	96.3	2025	6	AX327995	AX327995 Sequence
7	1759.6	96.3	2219	6	AX327993	AX327993 Sequence
8	1727	94.5	1956	9	HSA6701	AJ006701 Homo sapi
9	1330.4	72.8	3156	9	AK074411	AK074411 Homo sapi
10	851.4	46.6	2385	6	AX166526	AX166526 Sequence
11	851.4	46.6	2576	9	AB058714	AB058714 Homo sapi
12	851.4	46.6	2720	9	HSM805307	AL834275 Homo sapi
13	851.4	46.6	2897	6	AX642966	AX642966 Sequence
14	851.4	46.6	3007	9	AF479826	AF479826 Homo sapi
15	851.4	46.6	3109	9	AF479827	AF479827 Homo sapi
16	594	32.5	2333	9	AF020089	AF020089 Homo sapi
17	583.2	31.9	1873	9	BC024291	BC024291 Homo sapi
18	426	22.3	5609	3	AY060288	AY060288 Drosophil
19	391	21.4	2128	9	BC016681	BC016681 Homo sapi
20	355.4	19.5	3059	3	AF316542	AF316542 Caenorhab
21	349.2	19.1	3933	3	AB014885	AB014885 Halocynth
22	340	18.6	2123	3	AK116009	AK116009 Clona int
23	255	14.0	1791	9	AB062946	AB062946 Macaca fa
24	233	12.8	2259	10	AY151083	AY151083 Mus muscu
25	232.4	12.7	1594	6	AX056390	AX056390 Sequence
26	229.8	12.6	3170	6	AX305103	AX305103 Sequence
27	229.8	12.6	3250	6	AX305104	AX305104 Sequence
28	225	12.3	1549	6	AX685979	AX685979 Sequence
29	225	12.3	2462	6	AX399987	AX399987 Sequence
30	225	12.3	2650	9	AY120867	AY120867 Homo sapi
31	225	12.3	2954	6	BD127952	BD127952 Primer fo
32	225	12.3	2954	9	AK075372	AK075372 Homo sapi
33	225	12.3	3226	9	AK057448	AK057448 Homo sapi
34	225	12.3	3269	6	AX680149	AX680149 Sequence
35	225	12.3	3312	6	AX305105	AX305105 Sequence
36	225	12.3	3392	6	AX305106	AX305106 Sequence
37	225	12.3	3529	9	AB088047	AB088047 Homo sapi
38	225	12.3	3609	9	AB049127	AB049127 Homo sapi
39	225	12.3	4917	9	AB058763	AB058763 Homo sapi
40	224.6	12.3	2650	10	AB020480	AB020480 Rattus no
41	224.6	12.3	4092	10	AF106937	AF106937 Rattus no
42	222.8	12.2	2352	6	AX262512	AX262512 Sequence
43	222.8	12.2	2533	6	AX024729	AX024729 Sequence
44	222.8	12.2	2632	6	AX024732	AX024732 Sequence
45	222.8	12.2	2968	6	AX262510	AX262510 Sequence

ALIGNMENTS

RESULT 1
AX661193
LOCUS AX661193 1827 bp DNA linear PAT 22-MAR-2003
DEFINITION Sequence 3 from Patent WO02059287.
ACCESSION AX661193
VERSION AX661193.1 GI:29162845
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1
AUTHORS Turner C.A. and Mathur B.
TITLE Novel human kinases and polynucleotides encoding the same
JOURNAL Patent: WO 02059287-A 3 01-AUG-2002;

ORGANISM	Homo sapiens
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
AUTHORS	Turner,C.A. and Mathur,B.
TITLE	Novel human kinases and polynucleotides encoding the same
JOURNAL	Patent: WO 0205287-A 1 01-AUG-2002;
FEATURES	Lexicon Genetics Incorporated (US)
source	Location/Qualifiers
	1..2007
	/organism="Homo sapiens"
	/mol_type="genomic DNA"
	/db_xref="taxon:9606"
BASE COUNT	445 a 626 c 604 g 332 t
ORIGIN	
	Query Match 100.0%; Score 1827; DB 6; Length 2007;
	Best Local Similarity 100.0%; Pred. No. 0;
	Matches 1827; Conservative 0; Mismatches 0; Indels 0; Gaps 0
Qy	1 ATGAAGGTGGAGCGGGAGATCGCGATCTCTGAAGCTCATTTAGCACACCCACACGTCCTCTAAAG 60
Db	181 ATGAAGGTGGAGCGGGAGATCGCGATCTCTGAAGCTCATTTAGCACACCCACACGTCCTCTAAAG 240
Qy	61 CTGCACGACGTTTATGAAAAAATAATTTGTACTGTGCTAGAACACACGTGTCAAGT 120
Db	241 CTGCACGACGTTTATGAAAAAATAATTTGTACTGTGCTAGAACACACGTGTCAAGT 300
Qy	121 GGTGAGCTCTTCGACTACTCTGCTGGAAGGGAGGCTGACGCCCTAAGGAGGCTCGGAAG 180
Db	301 GGTGAGCTCTTCGACTACTCTGCTGGAAGGGAGGCTGACGCCCTAAGGAGGCTCGGAAG 360
Qy	181 TTCTTCCGGCAGATCATCTCTGCGCTGGACTTCTGCCACAGCAGCTTCCATATGCCACAGG 240
Db	361 TTCTTCCGGCAGATCATCTCTGCGCTGGACTTCTGCCACAGCAGCTTCCATATGCCACAGG 420
Qy	241 GATCTGAACCTGAAAACTCTCTGCTGGACAGAGAAACAATCGCATCGCAGACTTT 300
Db	421 GATCTGAACCTGAAAACTCTCTGCTGGACAGAGAAACAATCGCATCGCAGACTTT 480
Qy	301 GGCAATGGCGTCCCTCGAGGTTGGCAGACGCTGTGGAGACAGCTGTGGGTCCCGCCAC 360
Db	481 GGCAATGGCGTCCCTCGAGGTTGGCAGACGCTGTGGAGACAGCTGTGGGTCCCGCCAC 540
Qy	361 TACGCTCTCCCGAGTGATCCGGGGGAGAGTATGACGCGCGAAGCGCAGCTGTGG 420
Db	541 TACGCTCTCCCGAGTGATCCGGGGGAGAGTATGACGCGCGAAGCGCAGCTGTGG 600
Qy	421 AGCTGCGCGCTCATCTGTTGCGCTTGTGTTGGGGGCTTGCCTTTGACGATGACAA 480
Db	601 AGCTGCGCGCTCATCTGTTGCGCTTGTGTTGGGGGCTTGCCTTTGACGATGACAA 660
Qy	481 TTGGACAGCTGCTGGAGAGGTGAAGCGGGGGGTGTTCCCATGCGCGCATTTATCCGG 540
Db	661 TTGGACAGCTGCTGGAGAGGTGAAGCGGGGGGTGTTCCCATGCGCGCATTTATCCGG 720
Qy	541 CCCGACTGCCAGAGTCTGCTACGGGGCATGATCGAGGTGGAGCGCGCAGCGCGCTCACG 600
Db	721 CCCGACTGCCAGAGTCTGCTACGGGGCATGATCGAGGTGGAGCGCGCAGCGCGCTCACG 780
Qy	601 CTAGAGCACATTCAGAAAAACATATGGTATATAGGGGGCAAGATGAGCCCGAAACAGAG 660
Db	781 CTAGAGCACATTCAGAAAAACATATGGTATATAGGGGGCAAGATGAGCCCGAAACAGAG 840
Qy	661 CAGCCCATTTCTCGCAAGGTGCAGATTCGCTCGCTGCTGCCAGCGCTGGAGGACATCGACCCC 720
Db	841 CAGCCCATTTCTCGCAAGGTGCAGATTCGCTCGCTGCTGCCAGCGCTGGAGGACATCGACCCC 900
Qy	721 GACGTGCTGGACAGCATGCACCTACTTGGGCTGCTTCCGAGACCGCAACAAGCTGTGCGAG 780
Db	901 GACGTGCTGGACAGCATGCACCTACTTGGGCTGCTTCCGAGACCGCAACAAGCTGTGCGAG 960
Qy	781 GACCTGCTGTCGAGGAGGAGAACACAGGAGAGATGATTTACTTCTCTCTCTGACACCGG 840

AR232170 AR232170 2908 bp DNA linear PAT 20-DEC-2002
LOCUS
DEFINITION Sequence 1 from patent US 6455292.
ACCESSION AR232170
VERSION AR232170.1 GI:27274061
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 2908)
AUTHORS Shu, Y., Fan, W., Kovacs, K.F., Zidanic, M. and Jay, G.
TITLE Full-length serine protein kinase in brain and pancreas
JOURNAL Patent: US 6455292-A 1 24-SEP-2002;
FEATURES
Location/Qualifiers
source 1..2908
/organism="unknown"
BASE COUNT 603 a 948 c 879 g 478 t
ORIGIN
Query Match 99.9%; Score 1825.4; DB 6; Length 2908;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1826; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 ATGAAGTGGAGCGGAGATCGGATCCTGAAAGCTCATTTGAGCACCCCGACGTCCTAAAG 60
Db
286 ATGAAGTGGAGCGGAGATCGGATCCTGAAAGCTCATTTGAGCACCCCGACGTCCTAAAG 345
QY 61 CTGACGAGCTTTATGAAACAAAATAATTTGTACCTGGTCTAGAACACGTTGACGTT 120
Db
346 CTGACGAGCTTTATGAAACAAAATAATTTGTACCTGGTCTAGAACACGTTGACGTT 405
QY 121 GGTGAGCTCTTCGACTACCTGTGGAAGAGGGAGGCTGAGCCCTAAGGAGGCTCGGAAG 180
Db
406 GGTGAGCTCTTCGACTACCTGTGGAAGAGGGAGGCTGAGCCCTAAGGAGGCTCGGAAG 465
QY 181 TTCTTCGCGAGATCATCTCTCGCTGGACTTCTGCCACAGCACTTCCATATGCCACAGG 240
Db
466 TTCTTCGCGAGATCATCTCTCGCTGGACTTCTGCCACAGCACTTCCATATGCCACAGG 525
QY 241 GATCTGAACCTGAAACCTCTCTGTCGACGAGAGAACCAATCGCATCGCAGACTTT 300
Db
526 GATCTGAACCTGAAACCTCTCTGTCGACGAGAGAACCAATCGCATCGCAGACTTT 585
QY 301 GGCATGGGCTCCCTGCGAGCTGGGACAGCCCTGTTGGAGACAGCTGGGGTCCGCCAC 360
Db
586 GGCATGGGCTCCCTGCGAGCTGGGACAGCCCTGTTGGAGACAGCTGGGGTCCGCCAC 645
QY 361 TACGCTGCCCCGAGGTGATCCGGGGGAGAGATGACGGCCGGAAGCGGACGTTGG 420
Db
646 TACGCTGCCCCGAGGTGATCCGGGGGAGAGATGACGGCCGGAAGCGGACGTTGG 705
QY 421 AGCTGGGGGTGATCTCTGTTGCCCTGCTGGTGGGGGCTCTGCCCTTCAGCATGACAA 480
Db
706 AGCTGGGGGTGATCTCTGTTGCCCTGCTGGTGGGGGCTCTGCCCTTCAGCATGACAA 765
QY 481 TTGCGACAGCTGCTGGAGAGGTGAAGCGGGGCGTGTTCACATCGCCGACCTTTATCCCG 540
Db
766 TTGCGACAGCTGCTGGAGAGGTGAAGCGGGGCGTGTTCACATCGCCGACCTTTATCCCG 825
QY 541 CCCGATGCGCAGAGTCTGCTACGGGGCATGATCGAGGTGGACCGCGACGCGGCTCAGC 600
Db
826 CCCGATGCGCAGAGTCTGCTACGGGGCATGATCGAGGTGGACCGCGACGCGGCTCAGC 885
QY 601 CTAGAGCATCTCAGAAACACATATGGTATATAGGGGGCAGAAATGAGCCCGAACAGAG 660
Db
886 CTAGAGCATCTCAGAAACACATATGGTATATAGGGGGCAGAAATGAGCCCGAACAGAG 945
QY 661 CAGCCCATCTCTCGAAGGTGAGATCGGCTCGCTGCCAGCTCGGAGGACATCGACCCC 720
Db
946 CAGCCCATCTCTCGAAGGTGAGATCGGCTCGCTGCCAGCTCGGAGGACATCGACCCC 1005
QY 721 GACGTGTGGACAGCATGCACTCACTGGGCTCTTCGAGACCGCAACAAAGCTGTGCGAG 780
Db

1006 GACGTGCTGGACAGCATGCACTCACTGGGCTGCTTCGAGAGCCGCAACAAGTGTGTCAG 1065
QY 781 GACGTGCTGCTCCAGAGGAGAAACAGGAGAAAGATGATTTACTTCTCTCTGGACCGG 840
Db 1066 GACGTGCTGCTCCAGAGGAGAAACAGGAGAAAGATGATTTACTTCTCTCTGGACCGG 1125
QY 841 AAGAGAAAGGTACCCGAGCGAGGAGATGAGACCTGCCCGGAAACAGATAGACCTT 900
Db 1126 AAGAGAAAGGTACCCGAGCGAGGAGATGAGACCTGCCCGGAAACAGATAGACCTT 1185
QY 901 CCCCGAAGCGTGTGGACTCCCGGATCTGAAACCGGACGCGCAAGCGCGGCCAGAACGC 960
Db 1186 CCCCGAAGCGTGTGGACTCCCGGATCTGAAACCGGACGCGCAAGCGCGGCCAGAACGC 1245
QY 961 AATGTCATGAGGTGCTCAGGTGACGAGCGGCGCTCCCGGCTGCTGCGCGCGGGCC 1020
Db 1246 AAGTCTCATGAGGTGCTCAGGTGACGAGCGGCGCTCCCGGCTGCTGCGCGCGGGCC 1305
QY 1021 ATTGAGATGGCCAGCACGCGCAGAGGTCTCGGTTCATCAGCGGTGCTCTCTCAGGCTT 1080
Db 1306 ATTGAGATGGCCAGCACGCGCAGAGGTCTCGGTTCATCAGCGGTGCTCTCTCAGGCTT 1365
QY 1081 TCCACGACCCACTCAGCAGCCCCGGGTGACCCCTCACCCCTCACAAAGGGGAGTCCC 1140
Db 1366 TCCACGACCCACTCAGCAGCCCCGGGTGACCCCTCACCCCTCACAAAGGGGAGTCCC 1425
QY 1141 CTCCCGACCCCAAGGGGACACCTGTTCACACGCAAGAGAGAGCCCGGCTGGCACGGCC 1200
Db 1426 CTCCCGACCCCAAGGGGACACCTGTTCACACGCAAGAGAGAGCCCGGCTGGCACGGCC 1485
QY 1201 AACCCTCAGCGCCCGTCCAGCCCCAGCGTCCGAGGGGTGCCCTGGAGGGCGGCTCAAC 1260
Db 1486 AACCCTCAGCGCCCGTCCAGCCCCAGCGTCCGAGGGGTGCCCTGGAGGGCGGCTCAAC 1545
QY 1261 TCCATCAAGAAACAGCTTTCTGGGCTCACCCCGTTCACCGCCGGAACCTGCAAGTTCCG 1320
Db 1546 TCCATCAAGAAACAGCTTTCTGGGCTCACCCCGTTCACCGCCGGAACCTGCAAGTTCCG 1605
QY 1321 ACGCGGAGGAGATGTCACACCTGACACAGAGTCGTCGCCAGAGCTGGGGAAGTCC 1380
Db 1606 ACGCGGAGGAGATGTCACACCTGACACAGAGTCGTCGCCAGAGCTGGGGAAGTCC 1665
QY 1381 TGGTTTGGGAACCTTCATCAGCTGGAGAGGAGAGAGATCTTCGTGGTTCATCAAGAG 1440
Db 1666 TGGTTTGGGAACCTTCATCAGCTGGAGAGGAGAGAGATCTTCGTGGTTCATCAAGAG 1725
QY 1441 AAACCTCTGAGCTCCATCAAGGCTGACATCGTGAACGCTTCTCTGTCGATTCAGTCTC 1500
Db 1726 AAACCTCTGAGCTCCATCAAGGCTGACATCGTGAACGCTTCTCTGTCGATTCAGTCTC 1785
QY 1501 AGCCACAGCGTTCATCTCCCAACGAGCTTCGGGCGGAGTACAGGCCACGGGGGGGCA 1560
Db 1786 AGCCACAGCGTTCATCTCCCAACGAGCTTCGGGCGGAGTACAGGCCACGGGGGGGCA 1845
QY 1561 GCCGTGTTCCAGAGCGGTCAGTTCAGGTTCATATCATCTACACGAGGGTGGGGAG 1620
Db 1846 GCCGTGTTCCAGAGCGGTCAGTTCAGGTTCATATCATCTACACGAGGGTGGGGAG 1905
QY 1621 GCGCAGAGGAGAAACGGGATCTACTCCGTCACTTACCCCTGCTCTCAGGCCCCAGCGGT 1680
Db 1906 GCGCAGAGGAGAAACGGGATCTACTCCGTCACTTACCCCTGCTCTCAGGCCCCAGCGGT 1965
QY 1681 CGCTTCAAGAGGGTGGGAGACCATCCAGGCCGAGCTGCTGAGCACACAGACCCGCT 1740
Db 1966 CGCTTCAAGAGGGTGGGAGACCATCCAGGCCGAGCTGCTGAGCACACAGACCCGCT 2025
QY 1741 GCGGCCAGCAGCTTTGTGAGACACCACTTAATGTATGAAATGATGACGGGGCGCTTTC 1800
Db 2026 GCGGCCAGCAGCTTTGTGAGACACCACTTAATGTATGAAATGATGACGGGGCGCTTTC 2085
QY 1801 AATGTGGAATTTATCCCGAAAGTTAA 1827
Db 2086 AATGTGGAATTTATCCCGAAAGTTAA 2112

```
RESULT 4
AY166857
LOCUS
DEFINITION
AY166857
VERSION
KEYWORDS
SOURCE
ORGANISM
Homo sapiens
Homo sapiens (human)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 3516)
Direct Submission
Submitted (19-OCT-2002) School of Life Sciences, Fudan University,
Institute of Genetics, Handan Rd, 220, Shanghai 200433, China
Location/Qualifiers
1..3516
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/chromosome="11"
/tissue_type="brain tumor"
387..2393
/note="similar to CAD20436.1"
/codon_start=1
/evidence=not_experimental
/product="serine/threonine protein kinase isoform"
/protein_id="AA087839.1"
/db_xref="GI:27369415"
/translation="MTSTGKGGQAOYVGPYRLEKLGQGLVKLVGHVCTCK
VAIKTVNKKLSVESVLMKVEREIAITLKHHPVHLKLDVYENKKLYLVLEHVSGB
LFDVLVKKGRITPEAKRPFQIISALDPFCHSHSICHRLDLPENLLIDKNNIRIADF
GNASLQVDSLETSCGSPHYACPEVIRGEKIDGRKADVWSCGVILFALLVGLPFD
DNLKQLKVRGVFHPHFIPDPQSLRGMSEVDAARLTLLEHIQKHIWYIGKNE
PEPEQIIPRKVIISLPLEDIDPDLVMSHSLGCFRDRNKLQDLLESSEENQEMKY
FLLDLDRKERYSQEDEDLPNEIDPPRKRVDSPLNLRKRRPERKSMELSVTDGG
SPVPARRAIEMAGQORSRSISGASSGLSTSPASPRVTPHPSPRGSLPLTPKCTPVH
TPKSPACTPNPTPSSPSVGVGPWRARLINSIKNSFLGSFRPRKRLQVPTPEMSNL
TPESPRLAKKSWGNFISLEKEQIFVVIKDKPLSSIKADIHFAFUSIFLSLSHVSIS
QTSPELAKEYATGGAVFQKPVQDITYTEGGAQKENGIIYSTFTLLSGPSRRFR
VVEIQALLSTHPPRAQHLSDTTCNEMMTGRLSKGIIIPKS"
BASE COUNT 642 a 1135 c 1145 g 594 t
ORIGIN
Query Match 99.9%; Score 1825.4; DB 9; Length 3516;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1826; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 ATGAGGTGGAGCGGAGATCGGATCGTGGAGCTCATTGAGCACCCCGCCAGTCTTAAAG 60
DB 567 ATGAGGTGGAGCGGAGATCGGATCGTGGAGCTCATTGAGCACCCCGCCAGTCTTAAAG 626
QY 61 CTGACGACGCTTTATGAAAAAATAATTTGTACTGTGGTCTAGAACACAGTGTGAGGT 120
DB 627 CTGACGACGCTTTATGAAAAAATAATTTGTACTGTGGTCTAGAACACAGTGTGAGGT 686
QY 121 GGTGAGCTCTTCGACTACCTGTGTGAAGAGGGAGGCTGAGCCCTTAAGAGGCTCGGAAG 180
DB 687 GGTGAGCTCTTCGACTACCTGTGTGAAGAGGGAGGCTGAGCCCTTAAGAGGCTCGGAAG 746
QY 181 TTCTTCCGCGAGATCATCTCGCGCTGACCTTCGCCACAGCCACTCCATATGCCACAGG 240
DB 747 TTCTTCCGCGAGATCATCTCGCGCTGACCTTCGCCACAGCCACTCCATATGCCACAGG 806
QY 241 GATCTGAAACCTGAAACCTCTCTGCTGGACGAGAGAAACAAATCCGCATCGCAGACTTT 300
DB 807 GATCTGAAACCTGAAACCTCTCTGCTGGACGAGAGAACACATCCGCATCGCAGACTTT 866
QY 301 GGCATGGCGTCCCTGAGGTTGGGACAGCTGTGTTGGAGACAGCTGTGGGTCCCCCCAC 360
```

```

867 GGCATGGCGTCCCTGAGGTTGGCGACAGCGCTGTGGAGACCAAGCTGTGGGTCCCCCCAC 926
361 TAGCGCTGCCCGGAGGTATCCGGGGGAGAGATGATGACCGCGGAGGCGGAGCGTGTGG 420
927 TAGCGCTGCCCGGAGGTATCCGGGGGAGAGATGATGACCGCGGAGGCGGAGCGTGTGG 986
421 AGCTGCGCGCTCATCTCTGTTTCGCTCTGCTGGTGGGGCTCTGCCCTTCGACGATGACAAC 480
987 AGCTGCGCGCTCATCTCTGTTTCGCTCTGCTGGTGGGGCTCTGCCCTTCGACGATGACAAC 1046
481 TTGGGACAGCTGCTGGAGAGGTGAAGCGGGCGTGTTCACATGCGCGCATTTATCCCG 540
1047 TTGGGACAGCTGCTGGAGAGGTGAAGCGGGCGTGTTCACATGCGCGCATTTATCCCG 1106
541 CCGACTGCCAGAGTCTGTACCGGGCATGATCGAGTGGAGCGCGCACCGCGCTCAAG 600
1107 CCGACTGCCAGAGTCTGTACCGGGCATGAGGAGGTGGAGCGCGCACCGCGCTCAAG 1166
601 CTAGAGCACATTCAGAAACACATATATAGGGGCAAGATGAGCCCGAACACAGAG 660
1167 CTAGAGCACATTCAGAAACACATATATAGGGGCAAGATGAGCCCGAACACAGAG 1226
661 CAGCCCATTCCTCGAAGGTGCAGATCCGCTCGCTGCCAGCTTGAGGAGCATCGACCC 720
1227 CAGCCCATTCCTCGAAGGTGCAGATCCGCTCGCTGCCAGCTTGAGGAGCATCGACCC 1286
721 GAGTGTGTGACAGCATGCACTCACTGGGGTGTTCGAGACCGCACACAGCTGCTGCAG 780
1287 GAGTGTGTGACAGCATGCACTCACTGGGGTGTTCGAGACCGCACACAGCTGCTGCAG 1346
781 GACCTGTGTCCGAGGAGGAGAACACAGGAGAGATGATTTACTTCTCTCTCGAGACCGG 840
1347 GACCTGTGTCCGAGGAGGAGAACACAGGAGAGATGATTTACTTCTCTCTCGAGACCGG 1406
841 AAAGAAAGGTACCGGACCGACAGGAGATGAGAGCTGCGCCCCCGGAAACGAGATAGACCT 900
1407 AAAGAAAGGTACCGGACCGACAGGAGATGAGAGCTGCGCCCCCGGAAACGAGATAGACCT 1466
901 CCGCGAAGGTGTGGACTCCCGATGCTGAAACCGGACCGGACCGGCGCGGCGGAGACGC 960
1467 CCGCGAAGGTGTGGACTCCCGATGCTGAAACCGGACCGGACCGGCGGCGGAGACGC 1526
961 AAATCCATGGAGGTGCTCAGCTGACGAGCGGGGCTTCCCGGCTGCTGCGCGCGGCGCC 1020
1527 AAATCCATGGAGGTGCTCAGCTGACGAGCGGGCTTCCCGGCTGCTGCGCGCGGCGCC 1586
1021 ATTGAGATGGCCCGACGCGGACAGAGTCTCGGTTCATCAGCGGTGCTCTCTCAGGCTTT 1080
1587 ATTGAGATGGCCCGACGCGGACAGAGTCTCGGTTCATCAGCGGTGCTCTCTCAGGCTTT 1646
1081 TCACACAGCCACTCAGCAGCCCCCGGTGACCCCTCACCCCTCACCAAGGGCAGTCCC 1140
1547 TCACACAGCCACTCAGCAGCCCCCGGTGACCCCTCACCCCTCACCAAGGGCAGTCCC 1706
1141 CTCACACCCCCCAAGGGGACACCTGTCCACACGCAAGAGAGCGCGGCTGGCACGCC 1200
1707 CTCACACCCCCCAAGGGGACACCTGTCCACACGCAAGAGAGCGCGGCTGGCACGCC 1766
1201 AACCCACAGCCCCCGTTCAGCCCCCAGCGTGGAGGGGTGCCCTGGAGGGCGCGCTCAAC 1260
1767 AACCCACAGCCCCCGTTCAGCCCCCAGCGTGGAGGGGTGCCCTGGAGGGCGCGCTCAAC 1826
1261 TCACATCAAGAACAGCTTTCTGGGGTCAACCCGCTTCCACCGCGGAAACTGCAAGTTCCG 1320
1827 TCACATCAAGAACAGCTTTCTGGGGTCAACCCGCTTCCACCGCGGAAACTGCAAGTTCCG 1886
1321 ACGCCGAGGAGATGTCCAACTGACACAGAGTCTGCTCCCGAGAGCTGGCGAAGTCC 1380
1887 ACGCCGAGGAGATGTCCAACTGACACAGAGTCTGCTCCCGAGAGCTGGCGAAGTCC 1946
1381 TGGTTTGGGAACCTTCATCAGCTTGAGAGAGGAGAGATCTTCTGTGGTTCATCAAGAC 1440
```

Db 1947 TGCTTTGGGAACCTTCATCAGCCTGGAGAGGAGGAGCAGATCTTCGTGCTCATCAAGAC 2006
Qy 1441 AAACCTCTGAGCTCCATCAAGGCTGACATCGTGCGAGCCCTTCGTGCTGATTCCTCCAGTCTC 1500
Db 2007 AAACCTCTGAGCTCCATCAAGGCTGACATCGTGCGAGCCCTTCGTGCTGATTCCTCCAGTCTC 2066
Qy 1501 AGCCACAGCGTCTATCTCCCAACGAGCTTCGCGGCGAGGTACAAAGCCACGCGGGGGCCA 1560
Db 2067 AGCCACAGCGTCTATCTCCCAACGAGCTTCGCGGCGAGGTACAAAGCCACGCGGGGGCCA 2126
Qy 1561 GCCGTGTTCCAGAGCGGCTCAAGTTCCAGGTTGATATCACTTACACGAGAGGTGGGGAG 1620
Db 2127 GCCGTGTTCCAGAGCGGCTCAAGTTCCAGGTTGATATCACTTACACGAGAGGTGGGGAG 2186
Qy 1621 GCCAGAGGAGAACGGCATCTACTCCGTCACTTCACTGCTCTCAGGCCCCAGCGCT 1680
Db 2187 GCCAGAGGAGAACGGCATCTACTCCGTCACTTCACTGCTCTCAGGCCCCAGCGCT 2246
Qy 1681 CGCTTCAAGAGGCTGTGGAGACCATCCAGGCCAGCTGCTGAGCACACACGACCCGCT 1740
Db 2247 CGCTTCAAGAGGCTGTGGAGACCATCCAGGCCAGCTGCTGAGCACACACGACCCGCT 2306
Qy 1741 CGCGCCAGCACCTGTGAGACACCACTAACTGTATGGAATGATGACGGGGCGGCTTTCC 1800
Db 2307 CGCGCCAGCACCTGTGAGACACCACTAACTGTATGGAATGATGACGGGGCGGCTTTCC 2366
Qy 1801 AATGTGGATTATCCGAAAGTTAA 1827
Db 2367 AATGTGGATTATCCGAAAGTTAA 2393

RESULT 5
AR232171
LOCUS AR232171 3364 bp DNA linear PAT 20-DEC-2002
DEFINITION Sequence 3 from patent US 6455292.
ACCESSION AR232171
VERSION AR232171.1 GI:27274062
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 3364)
AUTHORS Shu, Y., Fan, W., Kovacs, K. F., Zidanic, M. and Jay, G.
TITLE Full-length serine protein kinase in brain and pancreas
JOURNAL Patent: US 6455292-A 3 24-SEP-2002;
FEATURES
source Location/Qualifiers
1..3364
/organism="unknown"
BASE COUNT 639 a 1078 c 1069 g 578 t
ORIGIN

Query Match 99.28; Score 1813; DB 6; Length 3364;
Best Local Similarity 99.84; Pred. No. 0;
Matches 1827; Conservative 0; Mismatches 0; Indels 4; Gaps 1;
Qy 1 ATGAAGTGGAGCGGGAGATCGCGATCCTCAAGCTCAATGAGCACCCCCACACGTCCTAAAG 60
Db 409 ATGAAGTGGAGCGGGAGATCGCGATCCTCAAGCTCAATGAGCACCCCCACACGTCCTAAAG 468
Qy 61 CTGCACGACGTTATGAAACAAATAATTT-----GTACTGGTGTAGAACACGTTGTC 116
Db 469 CTGCACGACGTTATGAAACAAATAATTTAGGTAGTCTGTGTAGAACACGTTGTC 528
Qy 117 AGTGGTGTAGCTTTCAGCTACCTGTGTGAAGAGGGGAGGCTCAGCGCTAAGGAGGCTCG 176
Db 529 AGTGGTGTAGCTTTCAGCTACCTGTGTGAAGAGGGGAGGCTCAGCGCTAAGGAGGCTCG 588
Qy 177 GAAGTTCTTCGGGAGATCATCTCTGGGCTGGACTTCTGCCACAGCCACTCCATATGCCA 236
Db 589 GAAGTTCTTCGGGAGATCATCTCTGGGCTGGACTTCTGCCACAGCCACTCCATATGCCA 648
Qy 237 CAGGATCTGAAACCTTGAAACCTTCGTGAGAGAGAGAACAACTCCGCTATCGCAGA 296

Db 649 CAGGATCTGAAACCTGAAACCTTCCTGCTGGACAGAGAAACAAACATCCGATCGCAGA 708
Qy 297 CTTTGGATATGCGCTCCCTGCAGGTTGGCGACAGCTGTTGGAGACCAAGCTGTGGGTCCCC 356
Db 709 CTTTGGATATGCGCTCCCTGCAGGTTGGCGACAGCTGTTGGAGACCAAGCTGTGGGTCCCC 768
Qy 357 CCACCTAGCCCTCCCGGAGGTGATCCCGGGGAGAGTATGACGCGCGGAGGCGGAGCT 416
Db 769 CCACCTAGCCCTCCCGGAGGTGATCCCGGGGAGAGTATGACGCGCGGAGGCGGAGCT 828
Qy 417 GTGAGCTGCGCGCTCATCTCTGTTGCTTCTGCTGGTGGGGCTCTGCCCTTCGACGATGA 476
Db 829 GTGAGCTGCGCGCTCATCTCTGTTGCTTCTGCTGGTGGGGCTCTGCCCTTCGACGATGA 888
Qy 477 CAACCTTGGACAGCTGCTGGAGAGGTGAAGCGGGGCTGTTCCACATGCGGCACTTTAT 536
Db 889 CAACCTTGGACAGCTGCTGGAGAGGTGAAGCGGGGCTGTTCCACATGCGGCACTTTAT 948
Qy 537 CCGGCCGACCTGCCAGAGCTCTGCTACCGGGCATGATGAGGTGACGCGCGCACGCGGCT 596
Db 949 CCGGCCGACCTGCCAGAGCTCTGCTACCGGGCATGATGAGGTGACGCGCGCACGCGGCT 1008
Qy 597 CACGCTAGAGCACATTCAGAAACACATATGTTATATAGGGGGCAAGATGAGCCCGAAC 656
Db 1009 CACGCTAGAGCACATTCAGAAACACATATGTTATATAGGGGGCAAGATGAGCCCGAAC 1068
Qy 657 AGACGAGCCCATCTCCGCAAGGTGCGAGATCCGCTCCGCTGCCAGCCCTGGAGGACATCGA 716
Db 1069 AGACGAGCCCATCTCCGCAAGGTGCGAGATCCGCTCCGCTGCCAGCCCTGGAGGACATCGA 1128
Qy 717 CCGGACGCTGCTGGACAGCATGCACTCACTGGGGCTGCTTCGAGACCGCAACAAGCTGCT 776
Db 1129 CCGGACGCTGCTGGACAGCATGCACTCACTGGGGCTGCTTCGAGACCGCAACAAGCTGCT 1188
Qy 777 GCAGACCTGCTGCTCGAGGAGGAGAACACGAGAGAGATGATTACTTCTCTCTCTGGA 836
Db 1189 GCAGACCTGCTGCTCGAGGAGGAGAACACGAGAGAGATGATTACTTCTCTCTCTGGA 1248
Qy 837 CCGGAAAGAAAGGTACCCGAGCGAGGAGATGAGACCTGCCCGCCCGGAAACGAGATAGA 896
Db 1249 CCGGAAAGAAAGGTACCCGAGCGAGGAGATGAGACCTGCCCGCCCGGAAACGAGATAGA 1308
Qy 897 CCTCTCCCGGAGCGTGTGGAATCCCGATGCTGAAACCGGCAACGCGAGCGCGCCAGA 956
Db 1309 CCTCTCCCGGAGCGTGTGGAATCCCGATGCTGAAACCGGCAACGCGAGCGCGCCAGA 1368
Qy 957 AGCAAAATCCATGAGGTGCTCAGCGTCAACGAGCGCGCTCCCGGCTCCCTGCGCGGG 1016
Db 1369 AGCAAAATCCATGAGGTGCTCAGCGTCAACGAGCGCGCTCCCGGCTCCCTGCGCGGG 1428
Qy 1017 GGCCATTGAGATGGCCCGAGCGCGAGAGGTCTCGGTCCATCAGCGGTGCTCTCTCAGG 1076
Db 1429 GGCCATTGAGATGGCCCGAGCGCGAGAGGTCTCGGTCCATCAGCGGTGCTCTCTCAGG 1488
Qy 1077 CTTTTCACACGACCACTCAGCAGCCCCCGGGTGAACCTTCAACCCCTCACCAGGGGCGAG 1136
Db 1489 CTTTTCACACGACCACTCAGCAGCCCCCGGGTGAACCTTCAACCCCTCACCAGGGGCGAG 1548
Qy 1137 TCCTCTCCCGACCCCAAGGGGAGACCTGTCCACGCGCAAGAGGAGCCCGGCTGGCAC 1196
Db 1549 TCCTCTCCCGACCCCAAGGGGAGACCTGTCCACGCGCAAGAGGAGCCCGGCTGGCAC 1608
Qy 1197 GCCCAACCCCGCCCGCTCCAGCGCCAGCGTCCGAGGGGTGCGCTCGAGGGCGCGCT 1256
Db 1609 GCCCAACCCCGCCCGCTCCAGCGCCAGCGTCCGAGGGGTGCGCTCGAGGGCGCGCT 1668
Qy 1257 CAACCTCCATCAAGAAACAGCTTTCTGGGTTCACCCCGCTTCCACCCCGGAAACTGCAAGT 1316
Db 1669 CAACCTCCATCAAGAAACAGCTTTCTGGGTTCACCCCGCTTCCACCCCGGAAACTGCAAGT 1728
Qy 1317 TCAGAGCGCGGAGGAGATGTTCCAACTGACACAGAGTCTGCTCCCGAGAGCTGGCGAAGAA 1376
Db 1729 TCAGAGCGCGGAGGAGATGTTCCAACTGACACAGTCTGCTCCCGAGAGCTGGCGAAGAA 1788

QY 1261 TCCATCAAGAACAGCTTTCTGGGCTACCCCGCTTCCACCGCGGAACCTGCAAGTTCCG 1320
Db 1441 TCCATCAAGAACAGCTTTCTGGGCTACCCCGCTTCCACCGCGGAACCTGCAAGTTCCG 1500
QY 1321 ACGCCGAGGAGATGTCACACCTGACACGAGTCTGTCCTCCAGAGTGGCGAAGAGTCC 1380
Db 1501 ACGCCGAGGAGATGTCACACCTGACACGAGTCTGTCCTCCAGAGTGGCGAAGAGTCC 1560
QY 1381 TGGTTTGGGAACTTCATCAGCTGGAGAGGAGGAGAGATCTTCGTGCTCATCAAGAC 1440
Db 1561 TGGTTTGGGAACTTCATCAGCTGGAGAGGAGGAGAGATCTTCGTGCTCATCAAGAC 1620
QY 1441 AAACCTCTGAGCTCCATCAAGGCTGACATCGTGACCGCTTCTGTCGATTCCTCCAGTCTC 1500
Db 1621 AAACCTCTGAGCTCCATCAAGGCTGACATCGTGACCGCTTCTGTCGATTCCTCCAGTCTC 1680
QY 1501 AGCCACAGGCTATCTCCCAACGAGTTCGCGGCCGAGTACAAAGCCACGGGGGGCCCA 1560
Db 1681 AGCCACAGGCTATCTCCCAACGAGTTCGCGGCCGAGTACAAAGCCACGGGGGGCCCA 1740
QY 1561 GCGGTGTTCCAGAGCGGCTCAAGTTCAGGTTGATATCACTACACGAGGAGTGGGGAG 1620
Db 1741 GCGGTGTTCCAGAGCGGCTCAAGTTCAGGTTGATATCACTACACGAGGAGTGGGGAG 1800
QY 1621 GCGCAGAGGAGAGCGGCTATCTCCGCTCAACCTTCAACCTGCTCTCAGGCCCCCAGCGCT 1680
Db 1801 GCGCAGAGGAGAGCGGCTATCTCCGCTCAACCTTCAACCTGCTCTCAGGCCCCCAGCGCT 1860
QY 1681 GCGTTCAAGAGGTTGGAGACCATCAGGCCCGAGTGTGAGCACACAGCCGCT 1740
Db 1861 GCGTTCAAGAGGTTGGAGACCATCAGGCCCGAGTGTGAGCACACAGCCGCT 1920
QY 1741 GCGGCCAGCAGCTTGTGAGACACAC 1766
Db 1921 GCGGCCAGCAGCTTGTGAGACACCCCC 1946

RESULT 7
AX327993
LOCUS AX327993 2219 bp DNA linear PAT 07-JAN-2002
DEFINITION Sequence 1 from Patent WO0181588.
ACCESSION AX327993
VERSION AX327993.1 GI:18098146
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
Meyers, R.
2246, protein kinase molecules and uses therefor
Patent: WO 0181588-A 1 01-NOV-2001;
Millennium Pharmaceuticals, Inc. (US)
FEATURES
source
1. .2219
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
53. .2077
/note="unnamed protein product"
/codon_start=1
/protein_id="CAD20436.1"
/db_xref="GI:18098147"
/translation="MTSTGDKGGAQAQYVGPYRLKTLGKGTGLVKLGAVHCVTCOK
VAIKVREKLSVLMKVEBIAIKLIEHPVHLKLDVYENKYLVLVLEHVSSE
LPDLVKKGRUTPREAKRFQRIISALDFCHSHICHRDLKPNENLLDERNRIADF
GNASLQVDSLSLETSCKGSPHYACVEVIRGEYDKRKADVSCGVLLFALLVGAIPFD
DMRLQLEKXRGVFNHPFTTPDCQSLRMSEVDAARLTLEHIQKHLYWYIGKNE
PEPEQPTPRKYQIRSLSDLEIDPDLDSMHSILGCFDRNKLQDLLESSENOEKMY
FLLDKREYPSQDEDLPPNEIDPPRKRVDSPLMRHKGRRPERKSMVLSVTDGG
SPVPARRALENAHQGRSRISIGASSGLSTSPVTPVTHPPSPRGSPLETPKGTVPV
TPKSPAGTNPFTPPSPSPVGRVRAKLSIKNSIFLGSFRRKRLQVTPPREMSNL
TPESPPELAKKSWFGNFISLEKEBQIFVVIKDKPLSLIKADIHAFILSLSHSVIS
..

OTSPAEYKATGGPAVFOKPVKFOVDITTYTEGGEAKENGIVSYFTTLLSGPSRRFRK
VVFETQALLSTHDPAPAHLSSEPPAPGLSWGAGLKGOKVATSYESSL"
BASE COUNT 476 a 716 c 683 g 344 t
ORIGIN
Query Match 96.3%; Score 1759.6; DB 6; Length 2219;
Best Local Similarity 99.8%; Pred. No. 3.7e-311;
Matches 1762; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
QY 1 ATGAAGTGGAGCGGAGATCGGATCTCTGAAGCTCATTTGAGCACCCCGACGCTCTAAAG 60
Db 233 ATGAAGTGGAGCGGAGATCGGATCTCTGAAGCTCATTTGAGCACCCCGACGCTCTAAAG 292
QY 61 CTGACGAGCTTTATGAAACAAATAATTTGTACTCTGCTAGAACACACGTTGTCAGGT 120
Db 293 CTGACGAGCTTTATGAAACAAATAATTTGTACTCTGCTAGAACACACGTTGTCAGGT 352
QY 121 GGTGAGCTCTTCGACTACCTGCTGGAAGAGGGAGGCTCACGCTTAAGGAGGCTCGGAAG 180
Db 353 GGTGAGCTCTTCGACTACCTGCTGGAAGAGGGAGGCTCACGCTTAAGGAGGCTCGGAAG 412
QY 181 TTCTTCGCGAGATCATCTCTGGCTGGACTTCTGCCACAGCCACTCCATATGCCACAG 240
Db 413 TTCTTCGCGAGATCATCTCTGGCTGGACTTCTGCCACAGCCACTCCATATGCCACAG 472
QY 241 GATCTGAAACCTGAAACCTCTCTGCTGGACGAGAAACAAACATCCGCTCGCAGACTTT 300
Db 473 GATCTGAAACCTGAAACCTCTCTGCTGGACGAGAAACAAACATCCGCTCGCAGACTTT 532
QY 301 GGCATGGCTCCCTGAGTTGGCGACACCTGTTGGAGACAGCTGTGGGTCCCCCAG 360
Db 533 GGCATGGCTCCCTGAGTTGGCGACACCTGTTGGAGACAGCTGTGGGTCCCCCAG 592
QY 361 TAGCCCTGCCCCGAGGTGATCCGGGGGAGAGTATGAGCGCGGAGCGGAGCGTGTGG 420
Db 593 TAGCCCTGCCCCGAGGTGATCCGGGGGAGAGTATGAGCGCGGAGCGGAGCGTGTGG 652
QY 421 AGCTGCGGCTCATCTCTGCTTGGCTGGGGGCTCTGCCCTTCGACAGATGACAAAC 480
Db 653 AGCTGCGGCTCATCTCTGCTTGGCTGGGGGCTCTGCCCTTCGACAGATGACAAAC 712
QY 481 TTGCGACAGCTGTGAGAAAGTGAAGCGGGGCTGTTTCACATGCGGCATTTATCCCG 540
Db 713 TTGCGACAGCTGTGAGAAAGTGAAGCGGGGCTGTTTCACATGCGGCATTTATCCCG 772
QY 541 CCCGACTGCCAGAGTCTGCTACGGGCGATGATGAGCGCGGAGCGCGCCGCTCACG 600
Db 773 CCCGACTGCCAGAGTCTGCTACGGGCGATGAGCGCGGAGCGCGCCGCTCACG 832
QY 601 CTAGAGCACATTTCAGAAACACATATGTTATAGGGGGCAAGATGAGCCCGAACCCAGAG 660
Db 833 CTAGAGCACATTTCAGAAACACATATGTTATAGGGGGCAAGATGAGCCCGAACCCAGAG 892
QY 661 CAGCCCATCTCTCGAAAGTGCAGATCCGCTCGCTCCAGCCCTGGAGGACATCGACCCC 720
Db 893 CAGCCCATCTCTCGAAAGTGCAGATCCGCTCGCTCCAGCCCTGGAGGACATCGACCCC 952
QY 721 GAGCTGCTGACAGCATGCACTCACCTGGGCTGCTCCGAGACCGCAACAGCTGTGTCAG 780
Db 953 GAGCTGCTGACAGCATGCACTCACCTGGGCTGCTCCGAGACCGCAACAGCTGTGTCAG 1012
QY 781 GACCTGCTGTCCGAGGAGGAGAACCCAGGAGAGATGATTTACTTCTCTCTCTGGACCGG 840
Db 1013 GACCTGCTGTCCGAGGAGGAGAACCCAGGAGAGATGATTTACTTCTCTCTCTGGACCGG 1072
QY 841 AAAGAAAGTACCCGAGCCAGGAGGATGAGGACTCGCCCCCGGAAACGAGATAGACCT 900
Db 1073 AAAGAAAGTACCCGAGCCAGGAGGATGAGGACTCGCCCCCGGAAACGAGATAGACCT 1132
QY 901 CCCCAGGAGGCTGTGACTCTCCCGATGCTGAACCGGACCGGCAAGCGGCGGCGGCAAGCGC 960
Db 1133 CCCCAGGAGGCTGTGACTCTCCCGATGCTGAACCGGACCGGCAAGCGGCGGCGGCGGCAAGCGC 1192

QY	961	AAATCCATGAGGTGCTCAGCGTGAAGGACGGCGGCTCCCGGTGCTCGCGGGCGGCC	1020
Db	1193	AAATCCATGAGGTGCTCAGCGTGAAGGACGGCGGCTCCCGGTGCTCGCGGGCGGCC	1252
QY	1021	ATTGAGATGCCACGACGCGGACAGAGGTCTCGGTCCATCAGCGGTGCTCCTCAGGCGCTT	1080
Db	1253	ATTGAGATGCCACGACGCGGACAGAGGTCTCGGTCCATCAGCGGTGCTCCTCAGGCGCTT	1312
QY	1081	TCCACAGCGCCACTCAGCAGCGCCCGGGTGACCCCTCACCCCTCACCAAGGGGCGAGTCCC	1140
Db	1313	TCCACAGCGCCACTCAGCAGCGCCCGGGTGACCCCTCACCCCTCACCAAGGGGCGAGTCCC	1372
QY	1141	CTCCACACCCCAAGGGGACACCTGTCCACAGCCAAAGGAGAGCCCGGCTGCGACGCC	1200
Db	1373	CTCCACACCCCAAGGGGACACCTGTCCACAGCCAAAGGAGAGCCCGGCTGCGACGCC	1432
QY	1201	AAACCCACGCGCCGCTCCAGCGCCAGCGTCCGAGGGGTGCGCTGAGGGCGGGCTCAAC	1260
Db	1433	AAACCCACGCGCCGCTCCAGCGCCAGCGTCCGAGGGGTGCGCTGAGGGCGGGCTCAAC	1492
QY	1261	TCCATCAGAAACAGCTTTCTGGGCTCACCCCGCTCCACCGCGGAAACTGCAAGTTCCG	1320
Db	1493	TCCATCAGAAACAGCTTTCTGGGCTCACCCCGCTCCACCGCGGAAACTGCAAGTTCCG	1552
QY	1321	ACGCGGAGAGATGTCACAACTGACACAGAGTCTGCCAGAGGTGCGGAGAGTCC	1380
Db	1553	ACGCGGAGAGATGTCACAACTGACACAGAGTCTGCCAGAGGTGCGGAGAGTCC	1612
QY	1381	TGGTTTGGAACTTATCAGCTGAGGAGGAGGAGCAGATCTTGTGTCTATCAAGAC	1440
Db	1613	TGGTTTGGAACTTATCAGCTGAGGAGGAGGAGCAGATCTTGTGTCTATCAAGAC	1672
QY	1441	AAACCTCTGAGCTCCATCAAGGCTGACATCTGTCAGCGCTTCTGTCGATTCCAGTCTC	1500
Db	1673	AAACCTCTGAGCTCCATCAAGGCTGACATCTGTCAGCGCTTCTGTCGATTCCAGTCTC	1732
QY	1501	AGCCACAGCGTCACTTCCCAAGAGCTTCCGGGCGGAGTACAAGGCCACGGGGGGCCA	1560
Db	1733	AGCCACAGCGTCACTTCCCAAGAGCTTCCGGGCGGAGTACAAGGCCACGGGGGGCCA	1792
QY	1561	GCGGTGTCAGAGCGGTCAAGTCCAGGTGATATCACCTACACGAGGAGGTGGGAG	1620
Db	1793	GCGGTGTCAGAGCGGTCAAGTCCAGGTGATATCACCTACACGAGGAGGTGGGAG	1852
QY	1621	GCGCAGAGGAGAAAGGATCTACTCCGTCACCTTCAACCTGCTCTCAGGCGCCAGCGT	1680
Db	1853	GCGCAGAGGAGAAAGGATCTACTCCGTCACCTTCAACCTGCTCTCAGGCGCCAGCGT	1912
QY	1681	CGCTTCAAGAGGTGTGAGACCATCCAGGCGGAGCTGCTGAGCACACAGACCGGCT	1740
Db	1913	CGCTTCAAGAGGTGTGAGACCATCCAGGCGGAGCTGCTGAGCACACAGACCGGCT	1792
QY	1741	GCGGCCACGACTTGTACAGACCCAC	1766
Db	1973	GCGGCCACGACTTGTACAGACCCCC	1998

RESULT 8
HSA6701
LOCUS HSA6701 1956 bp mRNA linear PRI 14-MAR-2001
DEFINITION Homo sapiens mRNA for putative serine/threonine protein kinase, partial.
ACCESSION AJ006701
VERSION AJ006701.1 GI:3217027
KEYWORDS putative; serine/threonine protein kinase.
SOURCE Homo sapiens
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1
AUTHORS Stanchi, F., Bertocco, E., Toppo, S., Dioguardi, R., Simonati, B., Cannata, N., Zimbelli, R., Lanfranchi, G. and Vallet, G.
TITLE Characterization of 16 novel human genes showing high similarity to

JOURNAL MEDLINE	Yeast sequences	Yeast 18 (1), 69-80 (2001)	21064499
PUBMED	2 (bases 1 to 1956)	11124703	
REFERENCE	Stanchi, F.		
AUTHORS	Submitted (02-JUN-1998)		
TITLE	Universita' di Padova, Via G. Colombo 3, 35121 Padova, 35121, ITALY		
JOURNAL	Location/Qualifiers		
FEATURES	1..1956		
source	/organism="Homo sapiens"		
	/mol_type="mRNA"		
	/db_xref="taxon:9606"		
	/sex="female"		
	/tissue="whole brain"		
	/clone_lib="Soares infant brain INIB"		
	/dev_stages="73 days post natal"		
	<1..1814		
CDS	/codon_start=3		
	/product="putative serine/threonine protein kinase"		
	/protein_id="CAA07196.1"		
	/db_xref="GI:3217028"		
	/translation="LIIHPHVLKLDVYENKKLYLVLEHVS GGELFDYLVKGRLLTP		
	KEARFPROIISALDFCHSHSICHRLKPKENLLDEKNIRIADFGMASLOVGSLLLE		
	TSCSPHYACPEVIRGEKYDGRKADVMSCGVILFALLVGLPFPDDNNLRQLLEKVRG		
	VFMHPHITPDQCSLKMSEVDAARLITLHIOHIMYIGGKNEPEPEPIPRKQI		
	RSLSPLSDIDPDLVDSHMSLGCDFRNLKLDLSEENQKMYIFLLDLRERYPSQ		
	DEDLPRNEIDPPRKVDSPMLNRHGRKRPERSMEVLSVTDGSPVPAARAEYMAQ		
	HQGRSRTISGASGLSPVPLSPRPVTPHPSGRGPLPTPGTVPHTKPESSPAGTNPMT		
	PPSPSGVGVPMRLANSIKNSFLGSPREFRRKLQVPTPEMSNLTPSSPELAKNSW		
	FGNFSLEKEQIIFVJLKDPLSLIKADIVHAFSLISPLSHSVISQTSFRAEYKATGG		
	FAVQKPVKFDVITYTEGGAQKNGIYVITFLLSGPSRRFRKRVETIQALLSTH		
	DPPAAQLHSEPPPPAPGLSMWAGLKGKQVATSYESSL"		
BASE COUNT	427 a 640 c 579 g 310 t		
ORIGIN			
Query Match	94.5%; Score 1727; DB 9; Length 1956;		
Best Local Similarity	99.7%; Pred. No. 3.3e-305;		
Matches 1730; Conservative	0; Mismatches 5; Indels 0; Gaps 0;		
QY	32 AGCTCATTTGAGCACCCCCACCGTCTCTAAAGCTGCACGAGCTTTATGAAAACAAAAATATT	91	
Db	1 AGCTCATTTGAGCACCCCCACCGTCTCTAAAGCTGCACGAGCTTTATGAAAACAAAAATATT	60	
QY	92 TGTACTGCTGCTAGAACACCGTGTGAGTGTGAGCTCTTCGACTACCTGTGTAAGAAGG	151	
Db	61 TGTACTGCTGCTAGAACACCGTGTGAGTGTGAGCTCTTCGACTACCTGTGTAAGAAGG	120	
QY	152 TGAGGCTGACCCCTTAAGGAGGCTCGAAGTCTTCCGGCAGATCATCTTCGCTGGACT	211	
Db	121 TGAGGCTGACCCCTTAAGGAGGCTCGAAGTCTTCCGGCAGATCATCTTCGCTGGACT	180	
QY	212 TCTGCCACAGCCACTTCCATATGCCACAGGATCTGAAACCTGAAACCTCTCTGCTGGAG	271	
Db	181 TCTGCCACAGCCACTTCCATATGCCACAGGATCTGAAACCTGAAACCTCTCTGCTGGAG	240	
QY	272 AGAGAAACAATCGGCATTCGACACTTTTGGATGCGCTTCCTCGAGTGTGCGACAGCC	331	
Db	241 AGAGAAACAATCGGCATTCGACACTTTTGGATGCGCTTCCTCGAGTGTGCGACAGCC	300	
QY	332 TGTGGACACGAGTGTGGTCCCGCCACTACGCTCGGCTCGGAGGTGATCCGGGGGAGA	391	
Db	301 TGTGGACACGAGTGTGGTCCCGCCACTACGCTCGGCTCGGAGGTGATCCGGGGGAGA	360	
QY	392 AGTATGACGGCCGGAAGGCGGACGTGTGGAGCTGCGGCGTTCATCTGTTGCTGGTGG	451	
Db	361 AGTATGACGGCCGGAAGGCGGACGTGTGGAGCTGCGGCGTTCATCTGTTGCTGGTGG	420	
QY	452 TGGGGGCTCTGCGCTTCGACGATGACAACTTGGACAGCTGCTGAGAGAGGTGAAGCGGG	511	
Db	421 TGGGGGCTCTGCGCTTCGACGATGACAACTTGGACAGCTGCTGAGAGAGGTGAAGCGGG	480	

QY 512 GCGTGTTCACATGCGCGCACTTTATCCGCGCGACTGCGCAGAGTCTGCTACCGGGGATGA 571
Db 481 GCGTGTTCACATGCGCGCACTTTATCCGCGCGACTGCGCAGAGTCTGCTACCGGGGATGA 540
QY 572 TCAGAGTGGACGCGCGCACTGCTACGCTAGAGCATTTCAGAAACACATATGTATA 631
Db 541 GCGAGTGGACGCGCGCACTGCTACGCTAGAGCATTTCAGAAACACATATGTATA 600
QY 632 TAGGGGCAAGATATGAGCCGAAACCGAGCAGGCCATTCTCCAAAGGTGAGATCCGCT 691
Db 601 TAGGGGCAAGATATGAGCCGAAACCGAGCAGGCCATTCTCCAAAGGTGAGATCCGCT 660
QY 692 CGTGCGCCAGCTGGAGGACATCGACCCCGACCTGCTGGACAGCATGCACTCACTGGCT 751
Db 661 CGTGCGCCAGCTGGAGGACATCGACCCCGACCTGCTGGACAGCATGCACTCACTGGCT 720
QY 752 GCTTCCGAGACCGCAACCAAGCTGCTGAGGACCTGCTGCCGAGGAGGAACCGAGAGA 811
Db 721 GCTTCCGAGACCGCAACCAAGCTGCTGAGGACCTGCTGCCGAGGAGGAACCGAGAGA 780
QY 812 AGATGATTACTTCTCTCTCTGAGACCGGAAGAAAGGTACCCGAGCCAGGAGATGAG 871
Db 781 AGATGATTACTTCTCTCTCTGAGACCGGAAGAAAGGTACCCGAGCCAGGAGATGAG 840
QY 872 ACCTGCCCCCGGAAACGAGATAGACCTCCCGGAGCGTGTGGACTCCCGATGCTGA 931
Db 841 ACCTGCCCCCGGAAACGAGATAGACCTCCCGGAGCGTGTGGACTCCCGATGCTGA 900
QY 932 ACCGGCAGCGCAAGCGCGCGCAGAAACCAAAATCCATGGAGGTGCTCAGCGTCAAGCGAG 991
Db 901 ACCGGCAGCGCAAGCGCGCGCAGAAACCAAAATCCATGGAGGTGCTCAGCGTCAAGCGAG 960
QY 992 GCGGCTCCCGGTGCTGCGCGCGGCGGCAATGAGATGGGCCAGACGCGCCAGAGTCTC 1051
Db 961 GCGGCTCCCGGTGCTGCGCGCGGCGGCAATGAGATGGGCCAGACGCGCCAGAGTCTC 1020
QY 1052 GGTCCATCAGCGGTGCTCTCAGGCGCTTCCACGACCCACCTCAGCAGCCCGCGGTGA 1111
Db 1021 GGTCCATCAGCGGTGCTCTCAGGCGCTTCCACGACCCACCTCAGCAGCCCGCGGTGA 1080
QY 1112 CCCCTCACCCCTCACCAAGGGGCGAGTCCCTTCCCTCCACCCCAAGGGGACACTGTCCACA 1171
Db 1081 CCCCTCACCCCTCACCAAGGGGCGAGTCCCTTCCCTCCACCCCAAGGGGACACTGTCCACA 1140
QY 1172 CGCCAAAGGAGAGCCCGCTGGACGCGCCCAACCCAGCCCGCTCCAGCCCGCAGGTGCG 1231
Db 1141 CGCCAAAGGAGAGCCCGCTGGACGCGCCCAACCCAGCCCGCTCCAGCCCGCAGGTGCG 1200
QY 1232 GAGGGGTGCTTGGAGGGCGGCTCAACTCCATCAAGAACAGCTTCTGGGTCAACCC 1291
Db 1201 GAGGGGTGCTTGGAGGGCGGCTCAACTCCATCAAGAACAGCTTCTGGGTCAACCC 1260
QY 1292 GCTTCCACCCCGGAAACTGCAAGTTCGACGCGCGGAGGAGATGTCCAACTGACACCCAG 1351
Db 1261 GCTTCCACCCCGGAAACTGCAAGTTCGACGCGCGGAGGAGATGTCCAACTGACACCCAG 1320
QY 1352 AGTGTCCCCAGAGCTGGGAGAGTCTGTGTTGGAACTTCATCAGCTTGGAGAGG 1411
Db 1321 AGTGTCCCCAGAGCTGGGAGAGTCTGTGTTGGAACTTCATCAGCTTGGAGAGG 1380
QY 1412 AGGAGCAGATCTTCTGGTTCATCAAGACAAACCTCTGAGCTCCATCAAGGCTGACATCG 1471
Db 1381 AGGAGCAGATCTTCTGGTTCATCAAGACAAACCTCTGAGCTCCATCAAGGCTGACATCG 1440
QY 1472 TGCAAGCTTCTCTGATTCCTAGTCTCAGCCACAGCGTCACTCCCAACAGAGTTC 1531
Db 1441 TGCAAGCTTCTCTGATTCCTAGTCTCAGCCACAGCGTCACTCCCAACAGAGTTC 1500
QY 1532 GGGCGGAGTACAGGCCACCGGGGGGCGAGCCGTGTTCCAGAGCCGCTCAAGTTCAGG 1591
Db 1501 GGGCGGAGTACAGGCCACCGGGGGGCGAGCCGTGTTCCAGAGCCGCTCAAGTTCAGG 1560

QY 1592 TTGATATCACCTACACGAGGGTGGGAGCGCGCAAGAGAGACCGCATCTACTCCGTCA 1651
Db 1561 TTGATATCACCTACACGAGGGTGGGAGCGCGCAAGAGAGACCGCATCTACTCCGTCA 1620
QY 1652 CTTTCAACCTGCTCTCAGCGCCCGCAGCGTTCGCTTCAAGAGGGTGGGAGACCATCCAGG 1711
Db 1621 CTTTCAACCTGCTCTCAGCGCCCGCAGCGTTCGCTTCAAGAGGGTGGGAGACCATCCAGG 1680
QY 1712 CCACGCTGCTAGGACACACGACCCCGCTGCGGCCAGCAGCTTGTTCAGACACAC 1766
Db 1681 CCACGCTGCTAGGACACACGACCCCGCTGCGGCCAGCAGCTTGTTCAGACACAC 1735

RESULT 9
AK074411 3156 bp mRNA linear PRI 15-FEB-2002
LOCUS Homo sapiens CDNA FLJ23931 fis, clone KARA00425, highly similar to
DEFINITION Homo sapiens mRNA for putative serine/threonine protein kinase.
ACCESSION AK074411
VERSION AK074411.1 GI:18677005
KEYWORDS oligo capping, fis (full insert sequence).
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1
AUTHORS Kawakami, T., Noguchi, S., Itoh, T., Shigeta, K., Senba, T.,
Matsumura, K., Nakajima, Y., Mizuno, T., Morinaga, M., Tanigami, A.,
Fujiwara, T., Ono, T., Yamada, K., Fujii, Y., Ozaki, K., Hirao, M.,
Ohmori, Y., Ota, T., Suzuki, Y., Obayashi, M., Nishi, T., Shibahara, T.,
Tanaka, T., Nakamura, Y., Isogai, T. and Sugano, S.
TITLE NEDO human cDNA sequencing project
JOURNAL Unpublished
2 (bases 1 to 3156)
AUTHORS Sugano, S., Suzuki, Y., Ota, T., Obayashi, M., Nishi, T., Isogai, T.,
Shibahara, T., Tanaka, T. and Nakamura, Y.
TITLE Direct Submission
JOURNAL Submitted (14-FEB-2002) Sumio Sugano, Institute of Medical Science,
University of Tokyo, Laboratory of Genome Structure, Human Genome
Center, Shirokane-dai, 4-6-1, Minato-ku, Tokyo 108-8639, Japan
(E-mail: cdna@ims.u-tokyo.ac.jp, Tel:81-3-5449-5286,
Fax:81-3-5449-5416)

COMMENT NEDO human cDNA sequencing project supported by Ministry of
Economy, Trade and Industry of Japan; cDNA full insert sequencing;
Research Association for Biotechnology; cDNA library construction,
5'- & 3'- end one pass sequencing; Department of Virology and Human
Genome Center, Institute of Medical Science, University of Tokyo
(partly supported by Science and Technology Agency).

FEATURES
source
1. .3156
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="KARA00425"
/tissue_type="ileal mucosa"
/clone_lib="KAIT"
/note="cloning vector: pME18SFL3"

BASE COUNT 588 a 1075 c 957 g 536 t
ORIGIN

Query Match 72.8%; Score 1330.4; DB 9; Length 3156;
Best Local Similarity 93.2%; Pred. No. 7.5e-233;
Matches 1442; Conservative 0; Mismatches 11; Indels 95; Gaps 1;

QY 219 CAGCCACTCATATGCCACAGGATCTGAAACCTGAAAACCTCTCTGACGAGAGAA 278
Db 3 CACCCTCTCTCCCGAGCCACAGGATCTGAAACCTGAAAACCTCTCTGACGAGAGAA 62
QY 279 CAACATCCGATCGCAGACTTTGGCATGGCTCCCTGCGAGTTGGCGACAGCTTTGGA 338
Db 63 CAACATCCGATCGCAGACTTTGGCATGGCTCCCTGCGAGTTGGCGACAGCTTTGGA 122
QY 339 GACCAGCTGTGGTCCCGCCCACTACGCTGCCCGCGAGGTGATCCCGGGGGAGAGTATGA 398

Db 123 GACCAAGCTGTGGGTCCCCCACTACGCTCCGCCGAGGTGATCCGGGGGAGAGTATGA 182
Qy 399 CGGCGGAGAGGCGGACGTGTGGAGCTGGCGGTTCATCTGTTCCGCTTGCTGTGGGGGC 458
Db 183 CGGCGGAGAGGCGGACGTGTGGAGCTGGCGGTTCATCTGTTCCGCTTGCTGTGGGGGC 242
Qy 459 TCTGCTCTTCGACATGATCAAACTTCGACAGCTGTGAGAGAGGTGAAGCGGGCGTGT 518
Db 243 TCTGCTCTTCGACG----- 256
Qy 519 CCACATGCGGCATTTTATCCCGCCCGACTGCCAGAGTCTGTACGGGGCATGATCGAGGT 578
Db 257 -----ATGATCGAGGT 267
Qy 579 GGAGCGCACGCGCCCTCAGCTAGACACATTCAGAAAAACATATGATATAGGGG 638
Db 268 GGAGCGCACGCGCCCTCAGCTAGACACATTCAGAAAAACATATGATATAGGGG 327
Qy 639 CAAGAATGAGCCCGAACACAGAGCAGCCCATTCCTCGCAAGGTGCAGATCCGCTGCTGCC 698
Db 328 CAAGAATGAGCCCGAACACAGAGCAGCCCATTCCTCGCAAGGTGCAGATCCGCTGCTGCC 387
Qy 699 CAGCTGAGGACATCGACCCCGACGCTGTGACAGCATGCACTACCTGCGGTGCTTCG 758
Db 388 CAGCTGAGGACATCGACCCCGACGCTGTGACAGCATGCACTACCTGCGGTGCTTCG 447
Qy 759 AGACCGCACAGCTGTGCGAGGACCTCTGTCGAGGAGGAGACACAGGAGAGATGAT 818
Db 448 AGACCGCACAGCTGTGCGAGGACCTCTGTCGAGGAGGAGAACACAGGAGAGATGAT 507
Qy 819 TTACTTCTCTCTCTGAGCGCGGAAAGATACCCGAGCGAGGAGATGAGGACCTGCC 878
Db 508 TTACTTCTCTCTCTGAGCGCGGAAAGATACCCGAGCGAGGAGATGAGGACCTGCC 567
Qy 879 CCCCAGGACGAGATAGACCTTCCCGGAGGAGTGTGACTTCCCGATGCTGAAACCGCA 938
Db 568 CCCCAGGACGAGATAGACCTTCCCGGAGGAGTGTGACTTCCCGATGCTGAAACCGCA 627
Qy 939 CGGCAAGCGGCGGCGAGAACCAATCATGAGGTGCTCAGCGTGAGCGGCGGCTC 998
Db 628 CGGCAAGCGGCGGCGAGAACCAATCATGAGGTGCTCAGCGTGAGCGGCGGCTC 687
Qy 999 CCGGTGCTGTGCGCGCGGCGCATTTGAGATGCGCCAGCAGCGCCAGAGGTCTCGGTCCAT 1058
Db 688 CCGGTGCTGTGCGCGCGGCGCATTTGAGATGCGCCAGCAGCGCCAGAGGTCTCGGTCCAT 747
Qy 1059 CAGCGGTGCTGTGCGCGCTTTTCCACGAGCCCACTCAGAGCGCCCGGGGTGACCCCTCA 1118
Db 748 CAGCGGTGCTGTGCGCGCTTTTCCACGAGCCCACTCAGAGCGCCCGGGGTGACCCCTCA 807
Qy 1119 CCGCTCACAGGGGCGAGTCCCTTCCCAACCCCAAGGGGACACCTGTCCACAGCCCAA 1178
Db 808 CCGCTCACAGGGGCGAGTCCCTTCCCAACCCCAAGGGGACACCTGTCCACAGCCCAA 867
Qy 1179 GGAGAGCGCGGTGGCGACGCCCAACCCAGCGCCCGCTCCAGCCCGAGCGGTGGAGGGGT 1238
Db 868 GGAGAGCGCGGTGGCGACGCCCAACCCAGCGCCCGCTCCAGCCCGAGCGGTGGAGGGGT 927
Qy 1239 GCGCTGAGGGGCGGCTCAACTCATCAAGAACAGCTTTCTGGGCTCAACCCCGTTCOA 1298
Db 928 GCGCTGAGGGGCGGCTCAACTCATCAAGAACAGCTTTCTGGGCTCAACCCCGTTCOA 987
Qy 1299 CCGCGGAACTGCAAGTTCCGACCGCGGAGAGATGTCCAACTTGACACAGAGTCTCTC 1358
Db 988 CCGCGGAACTGCAAGTTCCGACCGCGGAGAGATGTCCAACTTGACACAGAGTCTCTC 1047
Qy 1359 CCGAGAGCTGGGAGAGAGTCTGTTGGAACTTTCATCAGCTGGAGAGGAGGAGCA 1418
Db 1048 CCGAGAGCTGGGAGAGAGTCTGTTGGAACTTTCATCAGCTGGAGAGGAGGAGCA 1107
Qy 1419 GATCTTGTGTCATCAAGACAAACCTCTGAGCTTCATCAAGGCTGATCGTGACGC 1478
|||||

Db 1108 GATCTTCTGTGTATCAAGACAAACCTCTGAGCTCCATCAAGGCTGACATCGTGACGC 1167
Qy 1479 CTTCTCTGATTCCTCCAGTCTCAGCCACAGGTCATCTCCCAACAGAGCTTCCGGGCGGA 1538
Db 1168 CTTCTCTGATTCCTCCAGTCTCAGCCACAGGTCATCTCCCAACAGAGCTTCCGGGCGGA 1227
Qy 1539 GTACAAGCCACGCGGGGGCGAGCCGTGTTCCAGAACCGGTCAAGTTCACAGGTGATAT 1598
Db 1228 GTACAAGCCACGCGGGGGCGAGCCGTGTTCCAGAACCGGTCAAGTTCACAGGTGATAT 1287
Qy 1599 CACTACACGAGGTTGGGAGGCGAGAGGAGAACGAGCATCTACTCGTCACCTTCAC 1658
Db 1288 CACTACACGAGGTTGGGAGGCGAGAGGAGAACGAGCATCTACTCGTCACCTTCAC 1347
Qy 1659 CTTGCTCTCAGGCGCCAGCCGTCTTCAAGAGGTTGTGAGACCATCCAGGCGCCAGCT 1718
Db 1348 CTTGCTCTCAGGCGCCAGCCGTCTTCAAGAGGTTGTGAGACCATCCAGGCGCCAGCT 1407
Qy 1719 GCTGAGCACACACACCGCTGCGGCCAGCATTGTCAGACACCCAC 1766
Db 1408 GCTGAGCACACACACCGCTGCGGCCAGCATTGTCAGACACCCAC 1455
RESULT 10
AX166526 2385 bp DNA linear PAT*22-JUN-2001
LOCUS Sequence 17 from Patent WO0138503.
DEFINITION AX166526
ACCESSION AX166526
VERSION AX166526.1 GI:14546871
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1
AUTHORS Plowman,G.D., Whyte,D., Manning,G.S., Sudarsanam,S.S., Martinez,R.,
Flanagan,P. and Clary,D.S.
TITLE Novel human protein kinases and protein kinase-like enzymes
JOURNAL Patent: WO 0138503-A 17 31-MAY-2001;
Sugen, Inc. (US)
FEATURES
source
1. 2385
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
BASE COUNT 494 a 768 c 731 g 392 t
ORIGIN
Query Match 46.6%; Score 851.4; DB 6; Length 2385;
Best Local Similarity 68.9%; Pred. No. 2e-145;
Matches 1356; Conservative 0; Mismatches 416; Indels 195; Gaps 6;
Qy 1 ATGAAGTGGAGCGGAGATCGCGATCCTGAAGTCAATTGAGACACCCCGCTCTTAAG 60
Db 274 ATGAAGTGGAGCGGAGATCGCGATCCTGAAGTCAATTGAGACACCCCGCTCTTAAG 333
Qy 61 CTGCACGACGCTTATGAAAAACAAAAATATTTGACCTGTGTAGAACAGTGTCAAGT 120
Db 334 CTGCACGACGCTTATGAAAAACAAAAATATTTGACCTGTGTAGAACAGTGTCAAGT 393
Qy 121 GGTGAGCTTTTCGACTACCTGTTGAAGAGGGAGGCTGACGCTTAAGAGGCTCGGAG 180
Db 394 GGTGAGCTTTTCGACTACCTGTTGAAGAGGGAGAGTGTGACGCTTAAGAGGCTCGGAG 453
Qy 181 TTCTTCGGGAGATCATCTCTCGCTGGAATTCTGCCACAGCCCATTCATATGCCAGG 240
Db 454 TTCTTCGGGAGATCATCTCTCGCTGGAATTCTGCCACAGCTACTTCCATCTGCCACAGA 513
Qy 241 GATCTGAAACCTGAAAAACCTCTCTGACGAGAGAACAAACATCCGCTCGCAGACTTT 300
Db 514 GATCTGAAACCTGAAAAACCTCTCTTGGATGAGAAAAACAAACATCCGCTTGCAGACTTC 573
Qy 301 GGCATGGCGTCCCTGACAGTTGGCGACAGCCTGTTGGAGACCAAGTGTGGGTCCCCCAC 360

574 Db GGCATGGCGTCCCTGACAGTGGGGGACAGCTCTCTGGAGACAGCTGCGGGTCCCCCAT 633
361 QY TAGCCCTGCCCGGAGTATCGGGGGGAGAGTATGACGGCGGAGGCGGAGTGTGG 420
634 Db TAGCGGTCTCAGAGTGTATAGGGGGGAAATATATGATGGCGCGGCGAGACATGTGG 693
421 QY AGCTGCGGCTCATCTCTGTTGCGCTTCTGCTGGTGGGCTCTGCGCTTTCGACGATGACAA 480
694 Db AGCTGTGGAGTCTCTCTCTGCGCTCTCTGCTGGGGCTCTGCGCTTTCGATGACGACAA 753
481 QY TTGGCAGCTGTCTGAGAGAGTGTAGCGGGGCTGTTCACATGCGGCTTTATCCCG 540
754 Db CTCGCGCAGTGTCTGAGAGAGTGTAAACGGGGCTTTCACATGCGGCTTTATTCCT 813
541 QY CCGACTGCGCAGAGTCTGCTACGGGGGATGATCGAGGTGACCGCCGACCGCGCTCACG 600
814 Db CCAGATTGCCAGAGCTCTCTGAGGGGATGATCGAGTGTAGCGCGCGGAGGAGGCTCAGT 873
601 QY CTAGAGCACATTGAGAAACATATATGATATAGGGGGGAGAGTATGAGCCGAAAC 656
874 Db CTGGAGCAATTCAGAAACATCTCTGTTGTACTAGCGGGGAAACACAGAGCCGAGACCGCTGC 933
657 QY --AGAGAGCCCATCTCTGCGAGGTGACAGTCTGCTGCTGCTGCGGCTGAGGAGCATC 714
934 Db CTGGAGCCAGCCCTGCGCGCGGGTAGCCATGCGGAGCTGCGCATCCAAACGAGAGCTG 993
715 QY GACCCGACGCTGTGAGACAGCATCTCACTGCGGCTGCTTCCGAGACCGCAACAGAGCTG 774
994 Db GACCCGACGCTGTGAGAGCATGTCATCTGCGGCTGCTTCCGAGACCGCGAGAGCTG 1053
775 QY CTGACGAGCTGTGTCGAGGAGGAGAAACAGAGAGAGATGATTTACTTCTCTCTG 834
1054 Db CATCGAGCTGTGCGAGTGTGAGGAGGAGAAACAGAGAGAGATGATTTACTTCTCTG 1113
835 QY GACCGGAGAGAGTGTGAGGAGGAGGAGATGAGGAGGAGGAGGAGGAGGAGGAGGAG 894
1114 Db GATCGGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1173
895 QY GACCTCTCCGAGAGCTGTGAGTCTCCCGATGCTGAACCGGACCGGACCGGCGGCGCA 954
1174 Db GACCTCTCCGAGAGCTGTGAGTCTCCCGATGCTGAACCGGAGCGGAGCGGAGCA 1233
955 QY GAACGAAATTCATGAGGAGTGTCTCAGGTGACGGA-----CGCGGCTCCCGG 1005
1234 Db GAGCGGAATTCATGAGGAGTGTCTCAGGTGACGGA-----CGCGGCTCCCGG 1293
1006 QY CTTGCGCGGCGGCGCATTTGAGATGCGGCGGACGAGGAGTCTCGGTCCATCAGCGGT 1065
1294 Db CCGACCGAGCGGCGCTTTGGAGATGCGGCGGACGAGGAGTCTCGGTCCATCAGTGA 1353
1066 QY GCCTCTCAGGCGCTTTCCAGGAGGAGTCTCAGGAGGAGGAGGAGGAGGAGGAGGAG 1108
1354 Db GCCTCTCAGGCGCTTTCCAGGAGGAGTCTCAGGAGGAGGAGGAGGAGGAGGAGGAG 1413
1109 QY ----- 1108
1414 Db TCACCGGAGCGGCGGCTGTGAGATGAGGCTGAGGCGGCGGCGCTCCCGGAGTTCAGAAACG 1473
1109 QY -----TGACCCCTCACCCC 1122
1474 Db CAGAGCTGCTCTTCTGCGGCGGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1533
1123 QY TCACCAAGGGGAGTCTCCCTCTCC-----ACC 1149
1534 Db AGTGGCGCTCCACACCTCTGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1593
1150 QY CCCAAGGGGAGCCTGTGTCCACAGCGCAAGAGAGGAGGAGGAGGAGGAGGAGGAGGAG 1209
1594 Db CCGTGTGACTGCGCTCTGACACGCGCGCGGCGGAGTCCACCGGAGCGCGGAGCAACA 1653
1210 QY CCGCGTCCAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1263

1654 Db CCACCCCGAGCCCGCGGCTGCGGCGGAGCGCGCTGGAGGAGTGTCTCAACTCC 1713
1264 QY ATCAAGAAACAGCTTTCTGGGCTCACCCGCTTCCACCGCGGAAATGCAAGTTCGAGC 1323
1714 Db ATCCGCAACAGCTTCTGGGCTCCCTCGCTTTCACCGCGGCAAGATGCAAGTTCCTTACC 1773
1324 QY CCGAGGAGATGTCAAACCTGACACAGAGTGTCTCCAGAGCTGCGGAGAGTCTCTGG 1383
1774 Db GCTGAGGAGATGTCAAGCTTGAAGCGCAGAGTCTCTCCCGGAGCTGGCAAAACGCTCTCTGG 1833
1384 QY TTGGGAACTTCATCAAGCTGAGAGGAGGAGAGTCTCTCGTGTCTCAAGAGCAAA 1443
1834 Db TTGGGAACTTCATCTCTTGGACAAAGAGAAATATCTCTGCTGAAGAGCAAA 1893
1444 QY CTTCTGAGCTCATCAAGCTGACATCGTGTGACGCTTCTCTGATTTCCAGTCTCAGC 1503
1894 Db CTTCTCAGCAGCATCAAGCAGACATCGTCTCATGCTTCTCTGATCCAGCCTGAGT 1953
1504 QY CAGAGCTCATCTCCAAACAGAGCTTCCGGGCGGAGTACAAAGCCACCGGGGGCGCAGCC 1563
1954 Db CAGAGTGTCTCTCAGACACAGCTTTCAGGGCGGAGTACAAAGCCAGTGTGGCGGCGCTCC 2013
1564 QY GTTTCAGAAACCGCTCAAGTTCCAGGTTGATATCATCTACACGAGGCT----- 1614
2014 Db GTTTCAGAAACCGCTCGCTTCCAGGTGACATCATCTCTGAGGCTCCAGAGCC 2073
1615 QY -----GGGAGCGCAGAGGAGAACGATCTACTCTGCTCACTTCACTCTCATC 1665
2074 Db TCCCGCGACGCGCGGAGCGGAGTGTGTGATCTACTCTGCTCACTTCACTCTCATC 2133
1666 QY TCAGGCGCGGCGGCTTCAAGAGGAGTGTGAGAGCAGTCCAGGCGGCGGAGTGTGAGC 1725
2134 Db TCGGCTCCAGCGCTGCTTCAAGAGGAGTGTGAGAGCAGTCCAGGCGGAGTGTGAGC 2193
1726 QY ACACAGCGCCCGCTGCGGCGGCGGAGTGTGAGAGCAGTCCAGGCGGCGGAGTGTGAGC 1772
2194 Db ATCATGACGAGCGCTTCTGCGGCGGCGGAGTGTGAGAGCAGTCCAGGCGGAGTGTGAGC 2240

RESULT 11

AB058714
LOCUS AB058714 2576 bp mRNA linear PRI 05-JUN-2001
DEFINITION Homo sapiens mRNA for KIAA1811 protein, partial cds.
ACCESSION AB058714
VERSION AB058714.1 GI:14017838
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (sites)
AUTHORS Nagase,T., Nakayama,M., Nakajima,D., Kikuno,R. and Ohara,O.
TITLE Prediction of the coding sequences of unidentified human genes. XX. The complete sequences of 100 new cDNA clones from brain which code for large proteins in vitro
JOURNAL DNA Res. 8 (2), 85-95 (2001)
MEDLINE 21245130
PUBMED 11347906
REFERENCE 2 (bases 1 to 2576)
AUTHORS Ohara,O., Nagase,T. and Kikuno,R.
TITLE Direct Submission
JOURNAL Submitted (27-MAR-2001) Osamu Ohara, Kazusa DNA Research Institute, Department of Human Gene Research, 1532-3, Yana, Kisarazu, Chiba 292-0812, Japan (E-mail:cdnainfo@kazusa.or.jp, URL:http://www.kazusa.or.jp/huge, Tel:81-438-52-3913, Fax:81-438-52-3914)
FEATURES
source location/Qualifiers
1..2576 /organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="ba06731"
/tissue_type="brain"

gene	/note="vector:pBluescriptII SK plus"	
CDS	1..2576	
	/gene="KIAA1811"	
	<1..2149	
	/gene="KIAA1811"	
	/note="Start codon is not identified."	
	/codon_start=2	
	/product="KIAA1811 protein"	
	/protein_id="BAB47440.1"	
	/db_xref="GI:14017839"	
	/translation="TVNREKLSVLMKVEREIALKLIHHPHVLKLDVYENKYLX LVLEHVSGEPLDYLVKGRLLTPKEARFPFQIVSALDFCHSYISICHRDLPENLLD EKNIRIADFNASIQVDSLETSGSPHYACPEVIRKEKYDGRDRADWCGVILFA LLVGLALPDHEDDLRLLEKVRGVFMHPFIIPDQSLRLGMIEVEPEKRLSLQIQK HPWYLGGRHEDPCLEPAPGRVAMRSLPSNGELDPDVLWSMASLGCFRDRERLHREL RSEENQEKMIYLLLDKRYPCSDODLPNRNDVPRKRVDSPLMSLRGKRPER KSMVEVLSITDAGGSGSPVPTARALEMAOHSORSVSGASTGLASSPLSSPRSPFS SPEPAGDEARGGSPSTKQTLPSRGRGGAGGQPPPPARSTPLPFGPSRSG GTPHSPHPTASPTGPTGTPPPPGVGGVGAARSLNSIRNSFLGSPFRHRRKM QVPTAEHSSSLTPESPPELAKSWFNFISLDKEQIFLVLDKPLSSIKADIHVAEL SIPSLSHSVLSIPRAVYKASGSPVFPKRVQVDIISSEGEPPSPRRDGGGGI YSVFTLISGPRRRKRVVETIQALLSTHDOPSVOALADEKNGAQTRPAGAPPSLQ PPGPRDPELSSSPRGPPKXKKLLATNGTLP"	
BASE COUNT	508 a	842 c 766 g 460 t
ORIGIN		
Query Match	46.6%;	Score 851.4; DB 9; Length 2576;
Best Local Similarity	68.9%;	Pred. No. 2e-145;
Matches 1356; Conservative	0;	Mismatches 416; Indels 195; Gaps 6;
QY	1	ATGAAGGTGGAGCGGAGATCCGATCCCTGAAGCTCATTTGACGACCCCAAG 60
DB	38	ATGAAGGTGGAGCGGAGATCCCATCTCTGAAGCTCATCGAACACCCACATGCTCTCAAG 97
QY	61	CTGCACGAGTTTATGAACAAATAATTTGTACCTGGTCTAGACACGTTGTCAGT 120
DB	98	CTCCACGAGCTTACGAGAACAGAAATAATTTGTACCTGGTCTAGACACGTTGTCAGT 157
QY	121	GGTGAGCTCTCGACTTACCTGGTGAAGAGGGGAGGCTGACGCCCTAAGAGGCTCGGAAG 180
DB	158	GGTGAGCTTACGACTTACCTGGTGAAGAGGGGAGGCTGACGCCCTAAGAGGCTCGGAAG 217
QY	181	TTCTTCGGCAGATCATCTCGCTGGACTCTTGCACAGCCACTCCATATGCCACAGG 240
DB	218	TTCTTCGGCAGATTTGTCTCGCTGGACTCTTGCACAGCTACTCCATCTGCCACAGA 277
QY	241	GATCTGAACCTGAAACCTCTGCTGGACGAGAGAACAAATCGGATCCGACTTT 300
DB	278	GACCTAAAGCCCGAAGCTGCTTTTGGATGAGAAACAAATCCGCAATTGCAGACTTC 337
QY	301	GGCATGGGCTCCCTGCAGGTTGGCAGACAGCTGTTGGAGACAGCTGTGGGTCCCGCCAC 360
DB	338	GGCATGGGCTCCCTGCAGGTTGGGAGACAGCTTCTGGAGACAGCTGCGGTCCCGCCAT 397
QY	361	TACGCTGCGGAGTGATCGGGGGAGAGATGATGACGGCCGGAAGCGGAGCTGTGG 420
DB	398	TATGGGTGTCCAGAGTGATTAAGGGGAGAAATATGATGGCCGCGGAGACATGTGG 457
QY	421	AGCTGGGGCTCATCTGTTTCCCTTGTGTTGGGGCTCTGCCCTTCGACGATGACAAAC 480
DB	458	AGCTGTGGAGTTCATCTCTTCGCCCTTGTCTGGGGGGCTTGTGCCCTTTGATGACGACAAAC 517
QY	481	TTGCGACAGCTGCTGGAGAGGTGAAGCGGGGCGTGTTCACATGCGCGCACTTTATCCCG 540
DB	518	CTCCGCCAGCTGTGAGAGGTGAAACGGGGCGTCTTCCACATGCCCACTTCAATTCCT 577
QY	541	CCGACTGCGAGTGTGCTACGGGGCATGATCGAGGTGAGCGCGCAGCGGCTCTACG 600
DB	578	CCAGATTGCGAGGCTCTCTGAGGGGAATGATCGAAGTGGAGCCCGGAAAAAAGGCTCAGT 637
QY	601	CTAGAGCATTTAGAAACACATATGGTATATAGGGGCAAGATGAGCCCGAACCC--- 656
DB	638	CTGAGGCAAAATTCAGAAAATCTCTTGGTACCTAGGCGGGAAACACGAGCCAGACCCGTGC 697
QY	657	--AGAGCAGCCATTCTCGCAAGGTGCAGATCCGCTCGTCCAGCTCGAGCTTGGAGGACATC 714
DB	698	CTGGAGCAGCCCTTGGCCGCGGCTAGCCATGCGGAGCTGCCATCAACGGAGAGCTG 757
QY	715	GACCCCGAGTGTGGACAGATGACATCTACTGGGCTGCTTCCGAGACCGCAACAAGCTG 774
DB	758	GACCCCGAGCTCTTAGAGAGCATGGCATCACTGGGCTGCTTTCAGGACCGCGAGAGGCTG 817
QY	775	CTGCAGGACCTGCTGTCGAGGAGGAGAACACAGAGAGATGATTACTTCTCTCTCTG 834
DB	818	CATCGCAGCTGCCAGCTGAGGAGGAGAACCAAGAAAGATGATATATTTCTGCTTTTG 877
QY	835	GACCCGAAAGAAAGGTACCCGAGCAGAGGATGAGGACCTGCCCCCGCGAAACGAGATA 894
DB	878	GATCGGAAGGAGCGGTATCCAGCTGTGAGGACCAAGGACCTGCCCTCCCGGAATGATGTT 937
QY	895	GACCTCTCCCGGAAGCTGTGAGATCTCCCGATGTGAAACCGGACCGCAACGCGGCGCA 954
DB	938	GACCCCGCGGAGCGTGTGATTTCTCCATGCTGAGCCGTCACGGAAGCGCGACCA 997
QY	955	GAAACGCAATCCATGGAGGCTCTCAGCTGACGGA-----CGGCGGCTCCCGGTG 1005
DB	998	GAGCGAAGTCCATGGAAGTCTTGAGCATCACGATGCCGGGGTGTGGCTCCCTCTGTA 1057
QY	1006	CCTGCGCGCGGGCCATTGAGATGGCCAGCAGCGCCAGAGGTCTCGGTCTCATCAGCGGT 1065
DB	1058	CCACCCGAGCGGGCTTGGAGATGGCCAGCAGCAGAGATCCCGTAGGCTCAGTGA 1117
QY	1066	GCCTCTCAGGCTTTTCAACGAGCCCACTCAGCAGCCCGGG-----1108
DB	1118	GCCTCCAGGGTCTGTCTCCAGCCCTTAAGCAGCCCAAGGAGTCCGGTCTTTTCTCTTT 1177
QY	1109	-----1108
DB	1178	TCACCGGAGCGGGGGCTGGAGATGAGGCTCGAGGCGGGGCTCCCCGACTTCCAAACG 1237
QY	1109	-----TGACCCCTTCACCCC 1122
DB	1238	CAGACGCTGCTTCTCGGGGCCCCAGGGGTGGGGCGCGGGGAGCAGCCCCCGCCCC 1297
QY	1123	TCACCAAGGGGCACTCCCTCCCC-----ACC 1149
DB	1298	AGTCCCGCTCCACACCCCTGCGCGGCCCCCGAGGCTCCCGGCTCTCTCTGGGGGAGC 1357
QY	1150	CCCAAGGGGACACCTGTCCACACGCCAAAGAGAGCCGGCTGCGAGCCCAACCCACG 1209
DB	1358	CCCTTGCACTGCTCTGACACACGCCCGGGCCAGTCCCACCGGAGCCCCGGGAGCAACA 1417
QY	1210	CCCCCGTCCAGCCCC-----AGCGTGGGGGGTGCCTGAGGGGGCGGCTCAACTCC 1263
DB	1418	CCACCCCGAGCCCGGGGTGGCGTGGGGGAGCCGCTGGAGGAGTCTCTCAACTCC 1477
QY	1264	ATCAAGAACAGCTTTCTGGGCTCAACCCGCTTTCACCGCGGAACTGCAAGTTCGAGC 1323
DB	1478	ATCGCAACAGCTTCTGGGCTCCCTCGCTTTCACCGGCGCAAGATGCAAGTCCCTACC 1537
QY	1324	CCGAGGAGATGTCCAACTGACACAGAGTCTGCCAGAGCTGGGAGAGTCTCTGG 1383
DB	1538	GCTGAGGAGATGTCCAGTTGACGCCAGAGTCTCCCCGGAGCTGGCAAAACGCTCTCTGG 1597
QY	1384	TTTGGGAACCTTCATCAGCTGGAGAGGAGGAGAGATCTTCTGTTGTCATCAAGAGCAAA 1443
DB	1598	TTGCGGAACCTTCATCTCTTGGACAAAGAGAACAAATATTTCTCGTCTAAAGACAAA 1657
QY	1444	CCTCTGAGCTCCATCAAGGCTGACATCGTGACGCCCTTCTGTGATTCAGTTCAGTCTCAGC 1503
DB	1658	CCTCTGAGCAGCATCAAGAGCAGACATCGTCCATGCTTTCTGTGATCCCGGAGCTGAGT 1717
QY	1504	CACAGCTGTCTCCCAACGAGCTTCCGGGCGAGGTACAGGCCACGGGGGGGCGGAGCC 1563
DB	1718	CACAGTGTGTGTACAGAACGAGCTTCCAGGCCCGAGTACAGGCCAGTGGGGGCGCCCTCC 1777


```
QY 601 CTAGAGCACATTTCAGAAACACATATATGTTATATAGGGGCGCAAGATGAGCCCGAACC--- 656
Db 874 CTGAGGCAATTCAGAAACATCTCTGTTACTAGGCGGGAACACAGGACCGCGTGC 933
QY 657 --AGAGCAGCCCATTCCTCCAGAGGTGCAGATCCGTCGTCGCCAGCCTGGAGGACATC 714
Db 934 CTGAGCCAGCCCTCGGCCCGCGGTAGCCATCGGAGCTGCGATCCTCAACGAGAGCTG 993
QY 715 GACCCGAGCTGCTGGACATGCACTCACTGGGTCTTCGAGACCGCACAGCTG 774
Db 994 GACCCGAGCTCTAGAGAGCTGCGATCACTGGGTCTTCAGGAGACCGGAGGCTG 1053
QY 775 CTGAGGACCTGCTGTCGAGGAGGAGAACACGAGAGAGATGATTACTTCTCTCTCTG 834
Db 1054 CATCGGAGCTGCGCAGTGAGGAGGAGAACCAAGAAAGATATATATCTGCTTTG 1113
QY 835 GACCGGAAGAAAGTTACCGGACGAGGAGATGAGACCTGCCCGCCCGGAACGAGATA 894
Db 1114 GATCGGAAGAGGCGGTATCCCGCTGTGAGGACGAGACCTGCTCCCGGAATGATGTT 1173
QY 895 GACCTCCCGGAGCTGTGTGGATCCCGATGCTGAACCGGACGCGAGCGCGGCCA 954
Db 1174 GACCCCGCGGAGCGTGTGGATTCTCCCATGCTGAGCCGTACGCGGAGCGCGGACCA 1233
QY 955 GAACGCAATTCATGAGGTGCTCAGCGTGACGGA-----CGCGGCTCCCGGCTG 1005
Db 1234 GAGCGGAAGTCCATGGAAGTCTCTGAGCATCACCGATGCGCGGGGTGGTGGCTCCCTGTG 1293
QY 1006 CTGCGCGGGCGGCGCATGAGATGCGGCCAGACAGCGGACAGAGTCTCGGTCCATCAGCGGT 1065
Db 1294 CCCACCGGAGCGGCGCTTGGAGATGGGCCAGACAGCGGACAGAGATCCCGTAGCGTCAGTGA 1353
QY 1066 GCTCTCTCAGCGCTTCCACAGCGCCACTCAGACAGCGCCCGCGG----- 1108
Db 1354 GCCTTCAGCGGTCTGCTCTCAGCCCTCTAAGCAGCCCAAGGAGTCCGCTCTTCTCTT 1413
QY 1109 ----- 1108
Db 1414 TCACCGGAGCGGGGCTGGAGATGAGGCTCGAGCGGGGCTCCCGGACTTCCAAAACG 1473
QY 1109 -----TGACCCCTCACCCC 1122
Db 1474 CAGACGCTGCTTCTCGGGGCCCCAGGGGTGGGGGCGCGGGAGCAGCCCCCGCCCCC 1533
QY 1123 TCACCAAGGGCGAGTCCCTCCCTCC-----ACC 1149
Db 1534 AGTGCCCGCTCCACACCCCTGCGCGGCCCGCAGGCTCCCGCGCTCTCTGCGGGGACC 1593
QY 1150 CCCAAGGGGACACTGTGTCCACAGCCCAAGAGAGAGCCCGCTGCGACGCCCAACCCCAAG 1209
Db 1594 CCTTGTCACTCGCTCTGCAACAGCCCCCGGCGCAGTCCCAACCGGAGCCCCGGGACAA 1653
QY 1210 CCCCCGTCCAGCCCC-----AGGTCGAGGGGTGCTTGGAGGGCGCGCTCAACTCC 1263
Db 1654 CCACCCCCAGCCCCGCGGTGCGGTGCGGGGAGCCCGCTGGAGGAGTCTGTCTCAACTCC 1713
QY 1264 ATCAAGAACAGCTTCTGCGGCTCACCCCGCTTCCACCGCGGAACTGCAAGTTCGCAAG 1323
Db 1714 ATCGGCAACAGCTTCTGCGGCTCCCTCTGCTTTTCAACCGGCGCAAGATGCGAGTCCCTACC 1773
QY 1324 CCGGAGGAGATGTCCAACTGACACCAAGTGTGCTCCCGAGAGCTGGCGAAGAGTCTCTG 1383
Db 1774 GCTGAGGAGATGTCCAGCTTGACGCCAGAGTCTCTCCCGGAGCTGGCAAAACGCTCTCTG 1833
QY 1384 TTTGGGAACTTCATCGCTGAGAGGAGAGAGATCTCTGTTGCTCATCAAGACAAA 1443
Db 1834 TTCGGAACTTCATCTCTCTGGAAGAAAGAAACAAATATTCTCTGCTGAAGGACAAA 1893
QY 1444 CCTCTGAGCTCCATCAGGCTGATCTGTCAGCGCTTCTCTGTCAGCTCCAGTCTCAGC 1503
Db 1894 CCTCTCAGCAGCATCAAGAGAGACATGTCCTATGCTTCTGTCATGCTCCCGGCTGAGT 1953
```

```
QY 1504 CACAGCGTCATCTCCAAACAGAGCTTCGGGGCCGAGTACAAGGCCACCGGGGGCCAGCC 1563
Db 1954 CACAGTGTGCTGTACAGACCCAGCTTCAGGGCCGAGTACAAGGCCAGTGGCGGCCCTCC 2013
QY 1564 GGTTCAGAGAGCCCGGTCAAGTTCAGGTTGATATATCACTACACCGAGGCT----- 1614
Db 2014 GTCTTCCAAAGCCCGTCCGCTTCAGGTGGACATCAGCTCCTCTGAGGGTCCAGAGCCC 2073
QY 1615 -----GGGGAGCGCGAGAGGAGAAAGGCATCTACTCCGTCACTTCAACCTCTCTCTC 1665
Db 2074 TCCCGCGAGCGGACCGGAGGTTGTCATCTACTCCGTCACTTCACTCTCTCTCTCTC 2133
QY 1666 TCAGCGCCCGAGCCGCTCGCTTCAAGAGGTGTGGAGACATCCAGGCCAGCTGCTGAGC 1725
Db 2134 TCGGGTCCAGCGCTCGGTTCAAGCGAGTGTGGAGACCATCCAGGCACAGCTCTCTGAGC 2193
QY 1726 ACACAGCGCCCGCTCGGCGCCAGCACTTGTCCAGACACCACTAACTG 1772
Db 2194 ACTCATGACCGCCCTCGTGCAGGCCCTGGCAGACGAGAGAACGG 2240

RESULT 14
AF479826
LOCUS
DEFINITION
Homo sapiens putative serine/threonine protein kinase mRNA,
complete cds.
ACCESSION
AF479826
VERSION
AF479826.1 GI:19401870
KEYWORDS
SOURCE
Homo sapiens (human)
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1 (bases 1 to 3007)
AUTHORS
She X.Y., Guo J.H. and Yu L.
DIRECT SUBMISSION
Submitted (02-FEB-2002) School of Life Sciences, Laboratory of
Human Genes Research, Institute of Genetics, Fudan University, 220
Handan Road, Shanghai 200433, P. R. China
FEATURES
Location/Qualifiers
1..3007
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/chromosome="19"
/map="19q13"
/tissue_type="brain"
128..2512
/codon_start=1
/product="putative serine/threonine protein kinase"
/protein_id="AAL87697.1"
/db_xref="GI:19401871"
/translation="MVAGITLCKGPESPDVSVPERKDEVAGGGGEEAEERGRHA
QVGVRLKLTGKGTGLVKGIVHCITGKVAIKIVNREKLSVLMKVEREIAITLK
LIEHVLKLDHYENKKYLYLVLEHVSGLGFDLYLVKGRITPREARFPFQIVSAL
DFCHSYVLCRDLKPNLLDEKNRIADFGMSLQVDSLETSCGSPFVACPVV
KEGTDVGRADNMSCGVLIFALLGALPDDNLRQLLEKVKRGMVFMHPFIPDDCS
LURGMEVPEKRLSLEIQKHFWYLGKHEDPCLEPAPGRVAMRSLPSNGELDP
VLESMAISLCFRDRERLHRELSERENOKMIVYLLDRKERYPSCEDQDLPRNDVD
PRKQVDSPLSRHGKRPKRSMEVLSITDAGGGSPVPTRRALMAOHRSORSRVS
GASTGLSSPLSPSPVPSFSPGAGDEARGGSPSTQTQLPSRGRGAGAGQRP
PPPSARSTPLPGPPGSPSSGGTPLHSPHTPRASPTGTTPTPPSPGGVGGAAWR
SLNLSIRNSPLGSPRFRKMQVTAENSLTPSSPELAKRSNFGNFIISLDKSOI
FVLKDKPLSISIKADIVHAFSTPSLSHLSVLSQTPFRAEYKASGSPVFKPVPFVD
ISSSEGPSPRRDGGGGIYSVFTLLTSGPSRRFKRVVETIQALLSTHDDPSVOA
LADEKNGAQTTPAGAPPSPSLQPPRPDPDELSSPRRPPKDKLLATNGTLP"
BASE COUNT 636 a 940 c 906 g 525 t
ORIGIN
Query Match 46.6%; Score 851.4; DB 9; Length 3007;
Best Local Similarity 68.9%; Pred. No. 2e-145;
Matches 1356; Conservative 0; Mismatches 416; Indels 195; Gaps 6;
```

QY 1 ATGAGGTGGAGCGGAGATCGCATCTGTAAGAGCTCATTTGAGCACCCCCACGTCCTCTAAG 60
Db |||||
QY 401 ATGAGGTGGAGCGGAGATCGCATCTGTAAGAGCTCATTTGAGCACCCACATGTCCTCAAG 460
Db |||||
QY 61 CTGACAGAGCTTTATGAAAACAAAATAATTTGTACCTGTTGCTAGAACACAGTGTCTCAGGT 120
Db |||||
QY 461 CTCCACAGAGCTCTACGAGAACAGAAAATATTTGTACCTGTTGCTAGAACACAGTGTCTCAGGG 520
Db |||||
QY 121 GGTGAGCTCTTCGACTACCTGTTGAGAGAGGAGGCTGAGCCCTTAAGAGAGCTCGGAAG 180
Db |||||
QY 521 GGTGAGCTATTTCGACTACTGTTGAAGAGAGGAGACTGAGCCCAAGAGAGCCCGGAAG 580
Db |||||
QY 181 TTCTTCGGGAGATCATCTCTGCTGAGACTTTCTCCACAGCACTCCATATGCCACAGG 240
Db |||||
QY 581 TTCTTTCCGCGAATGTGTCTGCTGAGACTTTCTGCCACAGCTACTCCATCTGCCACAGA 640
Db |||||
QY 241 GATCTGAACCTGAAAACCTCTGCTGACGAGAGAGAACATCCGATCGCAGACTTT 300
Db |||||
QY 641 GACCTAAAGCCCGAGAACCTGCTTTTGTGATGAGAAAACAAACATCCGCAATTCGAGACTTC 700
Db |||||
QY 301 GGCATGGGCTCCCTGCAGGTTGGCGACAGCCCTGTTGGAGACAGCTGTGGGTCCCCCCAC 360
Db |||||
QY 701 GGCATGGGCTCCCTGCAGGTTGGCGAGACAGCCCTCTGGAGACAGCTGTGGGTCCCCCAT 760
Db |||||
QY 361 TAGCGCTGCCCGAGGTGATCGGGGGAGAAATGATGACGCGCGGAAGCGGAGCGTGTGG 420
Db |||||
QY 761 TATGCTGTGCAGAGGTGATTAAGGGGGAAAATATGATGGCGCGCGGCGAGACATGTGG 820
Db |||||
QY 421 AGCTCGGCGCTATCCTGTTGCGCTTGTGGTGGGGCTCTGCCCTTCGACGATGACAAAC 480
Db |||||
QY 821 AGCTGTGAGTATCCTCTTGCGCTGTCTGTTGGGGGCTCTGCCCTTTGTGATGACGACAAAC 880
Db |||||
QY 481 TTGCACAGCTGCTGAGAGGTGAAGCGGGCGGTGTTCCACATGCCGCACTTTATCCCG 540
Db |||||
QY 881 CTCGCCAGCTGCTGAGAGGTGAACCGGGCGTCTTCCACATGCCCCACTTCATTCT 940
Db |||||
QY 541 CCCGACTGCGAGCTGCTGACGAGCTGATCGAGGTGAGCGCGCACGCGGCTCTCAG 600
Db |||||
QY 941 CCAGATTGCCAGAGCTCTCTGAGGGGAATGATCGAAGTGGAGCCCGGAAAAGGCTCAGT 1000
Db |||||
QY 601 CTAGAGCATTTTCAAAACACATATGTTATATAGGGGCGAGAAATGAGCCGNAACC----- 656
Db |||||
QY 1001 CTGAGAGCAATTTCAAAACATCTTTGGTACTTAGCGGGGAAACAGAGCCAGACCCGTGC 1060
Db |||||
QY 657 --AGAGAGCCCATTTCTCGCAAGGTGCAGATCCGCTGCTGCCAGCGCTCGAGGACATC 714
Db |||||
QY 1061 CTGAGAGCCAGCCCTTGGCGCGGAGTACCATCGAGCGCTGCCATCAAAGGAGCTG 1120
Db |||||
QY 715 GACCCGAGCTGCTGAGCAGCATGCACTCATGCGGCTGCTTCGAGAGCCGCAACAGCTG 774
Db |||||
QY 1121 GACCCGAGCTCTTAGAGAGCATGTCATCATCTGGGCTGCTTCAGGGAACGCGAGAGGCTG 1180
Db |||||
QY 775 CTGAGGACCTGCTCTCGGAGGAGAGAACAGGAGAGAGATGATTTACTTCTCTCTCTG 834
Db |||||
QY 1181 CATCCGAGCTCGGAGTGTAGAGAGAGAACCAAGAAAAGATGATATATCTGCTTTTG 1240
Db |||||
QY 835 GACCGGAAGAAAGTATCCCGAGCGAGAGATGAGGACCTGCCCGCCCGGAACAGAGATA 894
Db |||||
QY 1241 GATCGGAAGGAGCGGTATCCCGAGCTGTGAGGACAGGACTCTGCCCTCCCGGAATGATGTT 1300
Db |||||
QY 895 GACCTCTCCCGGAAGCGGTGTGACTCCCGGATGCTGAACCGGACAGGCAAGCGGCGCA 954
Db |||||
QY 1301 GACCCCGCCCGGAAGCGGTGTGATTTCTCCCATGTCTGAGCGGTTCACGGGAAGCGGACCA 1360
Db |||||
QY 955 GAACCAATTCATGAGGTGCTCAGCGTGACGGA-----CGGCGGCTCCCGGTG 1005
Db |||||
QY 1361 GAGCGGAAGTCCATGGAAGTCTGAGCATCACCGATGCGGGGGTGTGGGTCTCCCTGTA 1420
Db |||||
QY 1006 CTGCGCGCGGGCCATTGAGATGCGCCAGCAGCGCCAGAGTCTCGGTCCATCAGCCGT 1065
Db |||||
QY 1421 CCCACCCGAGCGGCTTGTGAGATGCCCCAGCAGACAGAGATCCCGTAGGCTCAGTGA 1480
Db |||||
QY 1066 GCCTCTCAGGCTTTCCACACGAGCCCACTTCAGACAGCCCGCGG----- 1108

Db |||||
QY 1481 GCCTCCACGGGTGTGCTCTCCAGCCCTCTAAGCAGCCCAAGGAGTCGGTCTTTCTCTTT 1540
Db |||||
QY 1109 ----- 1108
Db |||||
QY 1541 TCACCGGAGCGCGGGGCTGGAGATGAGGCTCGAGCGGGGCTCCCGAGCTTCCAAAACG 1600
Db |||||
QY 1109 -----TGACCCCTTCACCCC 1122
Db |||||
QY 1601 CAGACGCTGCTTCTCGGGGCCCCAGGGGTGGGGGCCCGGGGAGCAGCCCGCCCGCCC 1660
Db |||||
QY 1123 TCACCAAGGGGAGTCCCTTCCCTCC-----ACC 1149
Db |||||
QY 1661 AGTCCCGCTCCACACCCCTGCGCGGCCCCAGGCTCCCGCGCTCTCTCTGGCGGAGC 1720
Db |||||
QY 1150 CCCAAGGGGACACTGTGTCACACGCAAGAGAGCCGCTGCGACGCCCAACCCACAG 1209
Db |||||
QY 1721 CCGTTGCACTCGCTCTGCAACAGCCCGGGGCCAGTCCACCGGGACCCCGGGGACAACA 1780
Db |||||
QY 1210 CCCCGCTCAGCCCC-----AGCGTCGAGGGGTGCCCTGGAGGGCGCGCTCAACTCC 1263
Db |||||
QY 1781 CCACCCCGAGCCCGGGGGTGGGTGGGGAGCCGCTGGAGGAGTGTCTCAACTCC 1840
Db |||||
QY 1264 ATCAAGAACAGCTTTCTGGGCTCAACCCGCTTCCACCGCGGAAACTGCAAGTTTCCAGC 1323
Db |||||
QY 1841 ATCGCAACAGCTTCTCTGGGCTCCCTCGCTTTCACCGCGCAAGATGCAAGTCCCTACC 1900
Db |||||
QY 1324 CCGAGAGAGATGTCACACTGACACAGAGTGTGTCGCCAGAGCTGGCGAAGAGTCTCTG 1383
Db |||||
QY 1901 GCTGAGAGATGTCAGCTTGAACGAGTGTCTCCCGAGCTGGCAAAACGCTCTCTG 1960
Db |||||
QY 1384 TTTGGGAACCTTCATCAGCTGAGAGAGAGCAGATCTTCGTGGTCTATCAAAAGACAAA 1443
Db |||||
QY 1961 TTGGGAACTTCATCTCTTGGACAAAGAGACAAATATTCCTGCTTAAAGACAAA 2020
Db |||||
QY 1444 CTTCTGAGCTCCATCAAGGCTGACATCTGACGCGCTTCTCTGTCGATTCCTCAGTCTCAGC 1503
Db |||||
QY 2021 CTTCTCAGCAGCATCAAGCAGACATCTCATCTCTTCTGTCGATCTCTGTCAGCTGAGT 2080
Db |||||
QY 1504 CACAGCTCATCTCCAAACAGCTTCCGGGCGGAGTACAGGCCACGGGGGGCGCAGCC 1563
Db |||||
QY 2081 CACAGTGTGCTCACAGACGAGCTTCAGGGCGGAGTACAGGCCAGTGGCGGCGCCCTCC 2140
Db |||||
QY 1564 GTGTTCCAGAGCCGCTCAAGTTCCAGCTTGAATATCACTACACGAGGGT----- 1614
Db |||||
QY 2141 GTCTTCCAAAAGCCGCTCCGCTTCCAGTGGACATCAGCTCTCTGAGGGTCCAGAGCCC 2200
Db |||||
QY 1615 -----GGGAGGCGCAGAAAGGAAACGGCATCTACTCCGTCACCTTCACTTCTGCTC 1665
Db |||||
QY 2201 TCCCGCGAGCGGACGCGAGGTGGTGGCATCTACTCCGTCACCTTCACTCTCATC 2260
Db |||||
QY 1666 TCAGGCGCCCGCTGCTTCAAGAGGTTGGTGGAGACCATCCAGGCCAGCTGCTGAGC 1725
Db |||||
QY 2261 TCGGGTCCCGAGCGCTGCTTCAAGCGAGTGGTGGAGACCATCCAGGCCAGCTCTGAGC 2320
Db |||||
QY 1726 ACACAGACCCGCTGCGGCGCCAGCACTTGTCAACACCACTAACTG 1772
Db |||||
QY 2321 ACTCATGACGAGCCCTCGTGCAGGCCCTGGCAGACGAGAAACGG 2367
Db |||||

RESULT 15
AF479827

LOCUS 3109 bp mRNA linear PRI 12-MAR-2002
DEFINITION Homo sapiens protein kinase-like protein mRNA, complete cds.

ACCESSION AF479827

VERSION AF479827.1 GI:19401873

KEYWORDS

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE 1 (bases 1 to 3109)

AUTHORS She,X.Y., Yu,L. and Guo,J.H.

Direct Submission
Submitted (02-FEB-2002) School of Life Sciences, Laboratory of
Human Genes Research, Institute of Genetics, Fudan University, 220
Handan Road, Shanghai 200433, P. R. China
Location/Qualifiers
1. 3109
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/chromosome="19"
278..2614
/codon_start=1
/product="protein kinase-like protein"
/protein_id="AAL87698.1"
/db_xref="GI:19401874"
/translation="MSSGAKEGGGSPAYHLPHPHPPHOAYVGPYRLKTLKGGQ
TGLVGLVHCITQGVAKIIVNREKLSVLMKVVEREIALKLIIEHPYRLKLDHYEN
KLLDLEKNIRIADFGNASLVQVDSLLSTSCGSPHYACPEVIGEKYIDRRADWSCG
VILFALVGNALPDDNLRQLLEKVRVMSLPNGELDPDVLSEWASLGCPRDRER
EQIKHPWYLGKHEPDPCEPAKRRVAMSLPSNGELDPDVLSEWASLGCPRDRER
LHRELSEENOEKMIYLLLDRLKRVPSCEDODLPRNDVDPRKRVDSPLMRHGK
RRERKSEVLSTADAGCGSPVTRALEMAHQSRSVSASGLSSPLSSPRS
PVFSFEPGADGARGGGSTKTQLPRGRPGGAGQPPPPARSRSTPLPGPGS
PRSGGTPLSPLHTPRASPTGTPTPPSPGGVGAARSPRSIRNSFLGSPRF
HRRMVPVTAEMSSLTPSSPELAKSWFNFTSLDKERQI FLVLKDKPLSLIKADI
VHAFLSPLSLHSVLSOTSFRAEYKASGSPVFOKPVRFVDISSEGEPPSPRRDGS
GGGIYSVTFTLISGSPRRKRVVETIQALLSTHODPSVQALADEKNGAQTRPAGAP
PRSLQPPPPRPDPELSSPRGPPKTKLLATNGTPLP"
BASE COUNT 611 a 1033 c 957 g 508 t
ORIGIN
Query Match 46.6%; Score 851.4; DB 9; Length 3109;
Best Local Similarity 68.9%; Pred. No. 2e-145;
Matches 1356; Conservative 0; Mismatches 416; Indels 195; Gaps 6;
QY 1 ATGAAGTGGAGCGGAGATCCGATCTGAAGCTCATTTGAGCACCCCGACGCTCTAAAG 60
DB ATGAAGTGGAGCGGAGATCCGATCTGAAGCTCATTTGAGCACCCCGACGCTCTCAAG 562
QY 61 CTGACGACGCTTATGAACAAATAATTTGACCTGGTGTAGACACGCTGCAGGT 120
DB CTGACGACGCTTATGAACAAATAATTTGACCTGGTGTAGACACGCTGCAGGT 622
QY 121 GGTGAGCTTCTCGACTACTCTGGTGAAGAGGGAGGCTGACGCTTAAGAGGCTCGGAAG 180
DB GGTGAGCTTCTCGACTACTCTGGTGAAGAGGGAGGCTGACGCTTAAGAGGCTCGGAAG 682
QY 181 TTCTTCCGCGAGATCATCTGCGCTGGACTTCTGCCACAGCTACTCTCCATCTGCCACAGA 240
DB TTCTTCCGCGAGATCATCTGCGCTGGACTTCTGCCACAGCTACTCTCCATCTGCCACAGA 742
QY 241 GATCTGAACCTGAAACCTCTGCTGACAGAGACACATCCGATCGCAGACTTT 300
DB GATCTGAACCTGAAACCTCTGCTGACAGAGACACATCCGATCGCAGACTTT 743
QY 301 GGCTATGGCGTCCCTGACGCTTGGCAGACCTGTTGGAGACAGCTGTGGGTCCCCCCCAC 360
DB GGCTATGGCGTCCCTGACGCTTGGCAGACCTGTTGGAGACAGCTGTGGGTCCCCCCCAC 802
QY 803 GGCTATGGCGTCCCTGACGCTTGGCAGACCTGTTGGAGACAGCTGTGGGTCCCCCCCAC 862
DB GGCTATGGCGTCCCTGACGCTTGGCAGACCTGTTGGAGACAGCTGTGGGTCCCCCCCAC 420
QY 361 TACGCTGCCCGGAGTGATCGGGGGAGAGATGATGACGCGCGGAGCGGAGCTGTGG 420
DB TACGCTGCCCGGAGTGATCGGGGGAGAGATGATGACGCGCGGAGCGGAGCTGTGG 922
QY 421 AGCTGCGCGCTCATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 480
DB AGCTGCGCGCTCATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 923
QY 481 TTGCGACAGCTGCTGGAGAGGTGAAGCGGGCGGTGTTTCCACATGCCGCACTTTATCCCG 540
DB TTGCGACAGCTGCTGGAGAGGTGAAGCGGGCGGTGTTTCCACATGCCGCACTTTATCCCG 982
QY 540 CTCCGCGAGCTGCTGGAGAGGTGAAGCGGGCGGTGTTTCCACATGCCGCACTTTATCCCG 1042
DB CTCCGCGAGCTGCTGGAGAGGTGAAGCGGGCGGTGTTTCCACATGCCGCACTTTATCCCG 1503
QY 541 CCCGACTGCCAGAGTCTGCTGACGGGGCATGATCGAGGTGGAGCGCGGCGCTCAGC 600

DB 1043 CCAGATTGCCAGAGCTCTCTGAGGGGAATGATCGAAGTGGAGCCCGAAAAAGGCTCAGT 1102
QY 601 CTAGAGCATTTCAGAAACACATATGTTATAGGGGGCAGAAATGAGCCGCAAC----- 656
DB 1103 CTGGAGCAAAATTCAGAAACATCTTTGGTACCTTAGGCGGGGAAACACGAGCCAGCCGTGC 1162
QY 657 --AGAGCAGCCCATTTCTCGCAAGGTGAGATCCCGCTCGCTGCCAGGCTGCGAGGACATC 714
DB 1163 CTGGAGCCAGCCCTTGGCGCGCGGTAGCCATGCGGAGCTGCCATCCAAACGAGAGCTG 1222
QY 715 GACCCCGAGCTGCTGGACAGCATGCACTCATCTGGGCTGCTTCCGAGACCGCAACAGCTG 774
DB 1223 GACCCCGAGCTGCTTAGAGAGCATGTCATCTGAGGCTGCTTCCAGGAGCCGCGAGAGCTG 1282
QY 775 CTGAGGAGCTGCTGTCGAGGAGGAGAACACAGGAGAGATGATTTACTTCTCTCTCTCTG 834
DB 1283 CATCGGAGCTGCGCAGTGGAGGAGAACACAGGAGAGATGATTTACTTCTCTCTCTCTG 1342
QY 835 GACCGGAAAGAAAGGTACCCGAGCCGAGGAGGATGAGGACCTGCCCCCGGAAACGAGATA 894
DB 1343 GATCGGAAAGGAGCGGTATCCAGCTGTGAGGACAGGACCTGCTCCCGGAATGATGTT 1402
QY 895 GACCTTCCCGGAAAGGTGTGACTCCCGATGCTGAAACCGGACAGGCAAGCGGCGGCCA 954
DB 1403 GACCCCGCGGAAAGCGTGTGATTTCTCCATGCTGAGCCGTACGCGGAAAGCGGCGACCA 1462
QY 955 GAACGCAAAATTCATGGAGGTGCTCAGCGTGACGGA-----CGCGGCTCCCCCGGTG 1005
DB 1463 GAGCGGAAGTCATGGAAGTCTCTGAGCATCACCGATGCCGGGGGTGTGGCTTCCCTCTGTA 1522
QY 1006 CTTGCGCGCGGGGCCATTGAGATGGCCAGCAGCGGACAGAGGTCTCGGTCTCATCAGCGGT 1065
DB 1523 CCCACCCGAGCGGCTTGGAGATGGCCAGCAGCAGAGATCCCGTAGCGTCACTGGA 1582
QY 1066 GCTTCTCTAGGCTTTTCCACGAGCCCATCTCAGAGCCCGCGG-----1108
DB 1583 GCCTCAGCGGTCTGCTCTCAGCGCTCTAAGCAGCCCAAGGAGTCCGCTCTTTCTCTTT 1642
QY 1109 -----1108
DB 1643 TCACCGGAGCGGGGGCTGAGATGAGGCTCGAGGGGGGGCTTCCCGACTTCCAAACG 1702
QY 1109 -----TGACCTCTACCC 1122
DB 1703 CAGACGCTGCTTCTCGGCGCCAGGGGTGGGGGCGCGCGGAGCAGCCCGCGCCCC 1762
QY 1123 TCACGAGGGGAGTCTCCCTTCCC-----ACC 1149
DB 1763 AGTGGCGGTCTCACACCTTCCCGCGCCCCCGCGCTTCCCGCGCTCTCTTGGCGGGACC 1822
QY 1150 CCCAAGGGGACACTCTGTCACACGCGCAAGGAGAGAGCCCGGCTGGCAGCGCCAAACCCACG 1209
DB 1823 CCTTGGACTGCTCTGTCACGCGCCCGGGCCAGTCCACCGGGACCCCGGGGACACA 1882
QY 1210 CCCCCGTCCAGCCCC-----AGCTGCGAGGGGTGCTTGGAGGGCGGGCTCAACTCC 1263
DB 1883 CCACCCCGCAGCCCGCGGTGGCGTGGGGAGCGGCTTGGAGGAGTCTCTCAACTCC 1942
QY 1264 ATCAAGAACAGCTTTCTGGGCTCACCCCGCTTCCAGCGCGGAACTGCAAGTTTCGAGC 1323
DB 1943 ATCCGCAACAGCTTCTTGGGCTCCCCCTCGCTTTCACCGGCGCAAGATGAGGTCCTTACC 2002
QY 1324 CCGGAGGAGATGTCCAACTTGACACAGAGTCTCTCCCGAGAGCTGCGGAGAGTCTCTGG 1383
DB 2003 GCTGAGGAGATGTCCAGCTTGACCGCAGAGTCTCTCCCGGAGCTTGGCAAAACGCTCTCTG 2062
QY 1384 TTTGGGAATCTTATCAGCTTGGAGAGGAGAGCAGATCTTCTGTGTCTATCAAGACAA 1443
DB 2063 TTCGGGAATCTTATCTCTTGGACAAAGAAAGAAATAATTCCTCGTGTAAAGGACAA 2122
QY 1444 CTTCTGAGCTTCTCAGGCTGACATGCTGCGAGCGCTTCTGTCGATTCCTGCTCTCAGC 1503

Db 2123 CCTCTCAGCAGCATCAAAGCAGACATCGTCCATGCTTTCTGTGATCCCCAGCCTGAGT 2182
QY 1504 CACAGCGTCATCTCCCAACAGAGCTTCGGGGCCGAGTACAAGGCCACGGGGGGCCAGCC 1563
Db 2183 CACAGTGTCTGTCAAGACCGAGCTTCAGGGCCGAGTCAAGGCCAGTGGCGGCCCTCC 2242
QY 1564 GTGTTCCAGAACCGGTCAAGTTCAGGTTGATATCACTACACGGAGGGT----- 1614
Db 2243 GTCTTCCAAAGCCGTCGGCTTCAGGTGGACATCAGCTCCTCTGAGGGTCCAGAGCCC 2302
QY 1615 -----GGGGAGGCGCAGAGGAGAACGGCATCTACTCCGTCACTTCACTCTGCTC 1665
Db 2303 TCCCCGCGACGGGACGGCAGCGGAGGTGGTGGCATCTACTCCGTCACTTCACTCTCATC 2362
QY 1666 TCAGGCCCCCAGCGCTCGCTTCAAGAGGGTGGTGGAGACCATCCAGGCCCCAGCTGCTGAGC 1725
Db 2363 TCGGGTCCCAGCGCTCGGTTCAAGCGAGTGGTGGAGACCATCCAGGCCACAGCTCCTGAGC 2422
QY 1726 ACACAGACCCCGCTGCGGCCCCAGCACTTGTCCAGACACCACTAACTG 1772
Db 2423 ACTCATGACCAAGCCCTCCGTGAGGCCCTGGCAGACGAGAGACGG 2469

Search completed: November 27, 2003, 09:04:15
Job time : 6558.39 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 27, 2003, 04:44:37 ; Search time 119.131 Seconds
(without alignments)
6769.052 Million cell updates/sec

Title: US-10-054-579-3
Perfect score: 1827

Sequence: 1 atgaagtgaggcgaggat.....gaattatccgaaagtaa 1827

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 569978 seqs, 220691566 residues

Total number of hits satisfying chosen parameters: 1139956

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents_NA.*

- 1: /cgn2_6/ptodata/2/ina/5A_COMB.seq.*
- 2: /cgn2_6/ptodata/2/ina/5B_COMB.seq.*
- 3: /cgn2_6/ptodata/2/ina/6A_COMB.seq.*
- 4: /cgn2_6/ptodata/2/ina/6B_COMB.seq.*
- 5: /cgn2_6/ptodata/2/ina/PCTUS_COMB.seq.*
- 6: /cgn2_6/ptodata/2/ina/backfiles1.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1825.4	99.9	2908	4	US-09-930-181-1
2	1813	99.2	3364	4	US-09-930-181-3
3	198.8	10.9	1742	3	US-08-557-006C-38
4	198.8	10.9	2652	3	US-08-557-006C-39
5	198.8	10.9	2761	3	US-08-557-006C-24
6	181.2	9.9	1647	3	US-09-101-146-44
7	177.4	9.7	2175	4	US-09-984-890-1
8	173.8	9.5	1747	3	US-08-557-006C-44
9	171.6	9.4	1929	4	US-09-359-161-4
10	168.4	9.2	2902	4	US-09-579-664B-4
11	151.6	8.3	1736	3	US-08-557-006C-37
12	151.6	8.3	1783	3	US-08-557-006C-36
13	146.4	8.0	2698	2	US-08-677-298-1
14	135.2	7.4	1429	2	US-09-159-385-4
15	135.2	7.4	1429	2	US-09-186-277-4
16	129.2	7.1	2132	3	US-09-159-385-3
17	129.2	7.1	2132	3	US-09-186-277-3
18	114	6.2	1599	3	US-09-256-465-1
19	114	6.2	1599	4	US-09-167-322-3
20	110.6	6.1	1694	4	US-09-579-664B-3
21	109.4	6.0	1480	4	US-09-016-434-1454
22	109	6.0	1733	4	US-09-620-312D-526
23	105	5.7	1507	4	US-09-523-849-4
24	105	5.7	1544	4	US-09-734-673-1
25	105	5.7	2060	4	US-09-523-849-1
26	101.8	5.6	1282	3	US-08-878-989-12
27	101.8	5.6	1282	3	US-09-272-796-12

28	101.8	5.6	1282	4	US-09-016-434-953	Sequence 953, Appl
29	101	5.5	2181	4	US-09-417-197-70	Sequence 70, Appl
30	101	5.5	2184	4	US-09-417-197-138	Sequence 138, Appl
31	101	5.5	2610	2	US-09-212-771-1	Sequence 1, Appl
32	101	5.5	2610	3	US-09-091-058-1	Sequence 1, Appl
33	100.8	5.5	3228	4	US-09-579-664B-5	Sequence 5, Appl
34	99.8	5.5	1104	4	US-09-691-861A-1	Sequence 1, Appl
35	99.8	5.5	9862	4	US-09-691-861A-3	Sequence 3, Appl
36	99.4	5.4	1051	4	US-09-523-849-3	Sequence 3, Appl
37	99.2	5.4	5228	4	US-09-428-711A-15	Sequence 15, Appl
38	98.2	5.4	1619	4	US-09-394-455-14	Sequence 14, Appl
39	96.4	5.3	3972	4	US-09-770-170-1	Sequence 1, Appl
40	95.8	5.2	488	4	US-09-523-849-24	Sequence 24, Appl
41	95.2	5.2	4935	2	US-08-631-097-3	Sequence 9, Appl
42	95.2	5.2	5886	3	US-08-810-712-9	Sequence 9, Appl
43	93.4	5.1	601	4	US-09-691-861A-14	Sequence 14, Appl
44	93	5.1	1732	4	US-09-430-564-1	Sequence 1, Appl
45	92.8	5.1	1788	4	US-09-417-197-68	Sequence 68, Appl

ALIGNMENTS

RESULT 1

US-09-930-181-1
; Sequence 1, Application US/09930181
; Patent No. 6455292

GENERAL INFORMATION:

; APPLICANT: Origene Technologies
; TITLE OF INVENTION: Full-Length Serine Protein Kinase in Brain and Pancreas
; FILE REFERENCE: 16U 101 V1
; CURRENT APPLICATION NUMBER: US/09/930,181
; CURRENT FILING DATE: 2001-08-16
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 1
; LENGTH: 2908
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (106)..(2112)
US-09-930-181-1

Query Match	99.9%	Score 1825.4;	DB 4;	Length 2908;
Best Local Similarity	99.9%	Pred. No. 0;		
Mismatches	0;	Indels	1;	Gaps 0;
Qy	1	ATGAAGTGGAGCGGAGATCGCGATCCTGAAGCTCATTGAGCAACCCCGCTCTAAAG	60	
Db	286	ATGAAGTGGAGCGGAGATCGCGATCCTGAAGCTCATTGAGCAACCCCGCTCTAAAG	345	
Qy	61	CTGCACACGCTTATGAAACAAAAATATTTGCTGCTAGACACGCTCTCAGGT	120	
Db	346	CTGCACACGCTTATGAAACAAAAATATTTGCTGCTAGACACGCTCTCAGGT	405	
Qy	121	GGTCAAGCTCTTCGACTACCTGTGTAAGAAGGGAGGCTGACGCTTAAGGAGCTCGGAAG	180	
Db	406	GGTCAAGCTCTTCGACTACCTGTGTAAGAAGGGAGGCTGACGCTTAAGGAGCTCGGAAG	465	
Qy	181	TTCTCCGGCAGATCATCTCTGCGCTGAGCTTTGCGCACAGCCACTCCATATGCCACAGG	240	
Db	466	TTCTCCGGCAGATCATCTCTGCGCTGAGCTTTGCGCACAGCCACTCCATATGCCACAGG	525	
Qy	241	GATCTGAAACCTGAAACCTCTCTGCTGAGAGAGAACACATCCGATCGCAGACTTT	300	
Db	526	GATCTGAAACCTGAAACCTCTCTGCTGAGAGAGAACACATCCGATCGCAGACTTT	585	
Qy	301	GGCATGCGCTCCCTGCGAGGTTGGGACAGCTGTGTGGAGACCACTGTGGGTCCCCCACC	360	
Db	586	GGCATGCGCTCCCTGCGAGGTTGGGACAGCTGTGTGGAGACCACTGTGGGTCCCCCACC	645	
Qy	361	TACGCTGCCCGGAGGTGATCCGGGGGAGAGATATGACGGCCGGAAGCGGACGTGTGG	420	

Db 646 TACGCTGCCCGGAGTGATCCGGGGGAGAGATGACGGCGGAGCGGAGCTGTGG 705
QY 421 AGCTGGGGGTGATCTCTGTTGGCTTGTGGTGGGGCTCTGCGCTTTCAGCATGACAAC 480
Db 706 AGCTGGGGGTGATCTCTGTTGGCTTGTGGTGGGGCTCTGCGCTTTCAGCATGACAAC 765
QY 481 TTCCGACAGCTGCTGGAGAAAGTGAAGCGGGGGTGTTCACATGCGGCACTTTATCCCG 540
Db 766 TTCCGACAGCTGCTGGAGAAAGTGAAGCGGGGGTGTTCACATGCGGCACTTTATCCCG 825
QY 541 CCGGACTGCCAGAGTCTGCTACGGGGCATGATCGAGGTGGAACCCGACGCGCTCAGG 600
Db 826 CCGGACTGCCAGAGTCTGCTACGGGGCATGATCGAGGTGGAACCCGACGCGCTCAGG 885
QY 601 CTAGAGCACATTCAGAAACACATATGGTATATAGGGGGGCAAGATGAGCCCGAACCAGAG 660
Db 886 CTAGAGCACATTCAGAAACACATATGGTATATAGGGGGGCAAGATGAGCCCGAACCAGAG 945
QY 661 CAGCCCATTCCTCGAAGGTGCAGATCCGCTCGCTGCCAGCTGGAGGACATCGACCCC 720
Db 946 CAGCCCATTCCTCGAAGGTGCAGATCCGCTCGCTGCCAGCTGGAGGACATCGACCCC 1005
QY 721 GAGTGTCTGACACATGCACTCACTGGGCTGCTTCCGAGACCCGCAACAGCTGCTGAG 780
Db 1006 GAGTGTCTGACACATGCACTCACTGGGCTGCTTCCGAGACCCGCAACAGCTGCTGAG 1065
QY 781 GAGTGTCTGCTCGAGGAGGAGAACACAGAGAGATGATTTACTTCTCTCTGGACCGG 840
Db 1066 GAGTGTCTGCTCGAGGAGGAGAACACAGAGAGATGATTTACTTCTCTCTGGACCGG 1125
QY 841 AAAGAAAGGTACCCGACGAGGAGATGAGACCTGCCCGGCGAACAGATAGACCT 900
Db 1126 AAAGAAAGGTACCCGACGAGGAGATGAGACCTGCCCGGCGAACAGATAGACCT 1185
QY 901 CCCCGBAAGCGTGTGAGTCTCCCGATGCTGAACCGGCAACGGCGAGCGCGGCGCAAGCG 960
Db 1186 CCCCGBAAGCGTGTGAGTCTCCCGATGCTGAACCGGCAACGGCGAGCGCGGCGCAAGCG 1245
QY 961 AATTCCATGGAGGTGCTCAGGTGACGAGCGGCGGCTCCCGGTGCTGCGCGCGGCGG 1020
Db 1246 AAGTGTCTGAGGAGTGTGCTCAGGTGACGAGCGGCGGCTCCCGGTGCTGCGCGCGGCG 1305
QY 1021 ATTGAGATGGCCCGACGAGCGGCTCGGTGCTCAGCGGTGCTCTCAGGCGCTT 1080
Db 1306 ATTGAGATGGCCCGACGAGCGGCTCGGTGCTCAGCGGTGCTCTCAGGCGCTT 1365
QY 1081 TCCACCGACCCCACTGAGCAGCCCGCGGTGACCTCTCACCCCTCAACAGGGGAGTCCC 1140
Db 1366 TCCACCGACCCCACTGAGCAGCCCGCGGTGACCTCTCACCCCTCAACAGGGGAGTCCC 1425
QY 1141 CTCCCCACCCCGACAGGAGACCTGTCCACACGCGCAAGGAGAGCGCGGTGCGACGCCC 1200
Db 1426 CTCCCCACCCCGACAGGAGACCTGTCCACACGCGCAAGGAGAGCGCGGTGCGACGCCC 1485
QY 1201 AACCCCACCGCCCGCTCCAGCGCCAGCGCTCGAGGGGGTGCCTCGAGGGGCGGCTCAAC 1260
Db 1486 AACCCCACCGCCCGCTCCAGCGCCAGCGCTCGAGGGGGTGCCTCGAGGGGCGGCTCAAC 1545
QY 1261 TCCATCAAGAAACAGTTTCTGGGTCAACCGCTTCCACCGCGGAACTCAAGTTCCG 1320
Db 1546 TCCATCAAGAAACAGTTTCTGGGTCAACCGCTTCCACCGCGGAACTCAAGTTCCG 1605
QY 1321 ACGCGGAGGAGATGCTCAACCTGACACGAGGTGCTCCCGAGAGCTGGCGAAGTCC 1380
Db 1606 ACGCGGAGGAGATGCTCAACCTGACACGAGGTGCTCCCGAGAGCTGGCGAAGTCC 1665
QY 1381 TGGTTTGGAACTTCATCAGCGCTGGAGAGGAGGAGATCTTCGTGGTTCATCAAGAG 1440
Db 1666 TGGTTTGGAACTTCATCAGCGCTGGAGAGGAGGAGATCTTCGTGGTTCATCAAGAG 1725
QY 1441 AAACCTCTGAGTTCATCAAGGCTGACATCTGTGCAACGCTTCTGTGCTGATTCAGTCTC 1500

Db 1726 AAACCTCTGAGTCCATCAAGCTGACATCGTGACGCTTCTGTCTGATTCCTCCAGTCTC 1785
QY 1501 AGCCACAGCTCATCTCCAAACGAGCTTCCGGGCGGAGTACAGGCCACGGGGGGGCA 1560
Db 1786 AGCCACAGCTCATCTCCAAACGAGCTTCCGGGCGGAGTACAGGCCACGGGGGGGCA 1845
QY 1561 GCGGTGTTCCAGAAACCGGTCAAGTTCCAGGTGATATCACTACACGAGGGTGGGAG 1620
Db 1846 GCGGTGTTCCAGAAACCGGTCAAGTTCCAGGTGATATCACTACACGAGGGTGGGAG 1905
QY 1621 GCGGAGAGGAGAACGGCATCTACTCGTCACTTCACTTCACTTCACTTCACTTCACTT 1680
Db 1906 GCGGAGAGGAGAACGGCATCTACTCGTCACTTCACTTCACTTCACTTCACTTCACTT 1965
QY 1681 CCGTTCAGAGGGTGTGGAGACCATCCAGGCCAGCTGCTGAGCACACACACCGGCT 1740
Db 1966 CCGTTCAGAGGGTGTGGAGACCATCCAGGCCAGCTGCTGAGCACACACACCGGCT 2025
QY 1741 GCGGCCAGCAGCTTGTGAGACACCACTAACTGTATGGAATGATGACGGGGCGGCTTCC 1800
Db 2026 GCGGCCAGCAGCTTGTGAGACACCACTAACTGTATGGAATGATGACGGGGCGGCTTCC 2085
QY 1801 AAATGTGGAATTCCTCCGAAAGTTAA 1827
Db 2086 AAATGTGGAATTCCTCCGAAAGTTAA 2112

RESULT 2

US-09-930-181-3
; Sequence 3, Application US/09930181
; Patent No. 6455292
; GENERAL INFORMATION:
; APPLICANT: OriGene Technologies
; TITLE OF INVENTION: Full-length Serine Protein Kinase in Brain and Pancreas
; FILE REFERENCE: 16U 101 V1
; CURRENT APPLICATION NUMBER: US/09/930,181
; CURRENT FILING DATE: 2001-08-16
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 3
; LENGTH: 3364
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (482)..(2239)
US-09-930-181-3

Query Match 99.2%; Score 1813; DB 4; Length 3364;

Best Local Similarity 99.8%; Pred. No. 0;

Matches 1827; Conservative 0; Mismatches 0; Indels 4; Gaps 1;

QY 1 ATGAAGGTGGAGCGGAGATCGGATCCCTGAGCTCATTGAGCACCCCGCTCTTAAG 60
Db 409 ATGAAGGTGGAGCGGAGATCGGATCCCTGAGCTCATTGAGCACCCCGCTCTTAAG 468
QY 61 CTGACAGACGCTTTATCAAAACAAAAATATTT---GTACCTGGTGTAGAAACGCTGTC 116
Db 469 CTGACAGAGCTTTATGAAACAAAAATATTTGTAGGTACCTGGTGTAGAAACGCTGTC 528
QY 117 AGGTGTGAGCTTTCGATPACTGTGTGAAGAGGGAGGCTGACCGCTTAAGAGGCTCG 176
Db 529 AGGTGTGAGCTTTCGATPACTGTGTGAAGAGGGAGGCTGACCGCTTAAGAGGCTCG 588
QY 177 GAAGTTCTTCCGCGAGATCATCTCGGCTGGAATTTCTGCCACAGCACTCCATATGCCA 236
Db 589 GAAGTTCTTCCGCGAGATCATCTCGGCTGGAATTTCTGCCACAGCACTCCATATGCCA 648
QY 237 CAGGGATCTGAACCTGAAACCTCTCTGTGTGAAGAGGAAACAAATCCGCTATCGCAGA 296
Db 649 CAGGGATCTGAACCTGAAACCTCTCTGTGTGAAGAGGAAACAAATCCGCTATCGCAGA 708
QY 297 CTTTGGCATGGCGTCCCTGCTGAGGTTGGGAGACGCTGTTGGAGACGAGCTGTGGTCCCC 356

Db 709 CTTTGGCATGGCGTCCCTGCAAGTTGGCGACAGACCTGTTGGAGACAGCTGTGGGTCCCC 768
QY 357 CCACTACGCTCCCGGAGGTGATCCGGGGGAGAGTATGACGCCCGGAGGGCGGAGCT 416
Db 769 CCACTACGCTCCCGGAGGTGATCCGGGGGAGAGTATGACGCCCGGAGGGCGGAGCT 828
QY 417 GTGAGCTGCGGCGTCACTCTGTTGCGCTTGTGCTGGGGGCTCTGCGCTTCGACGATGA 476
Db 829 GTGAGCTGCGGCGTCACTCTGTTGCGCTTGTGCTGGGGGCTCTGCGCTTCGACGATGA 888
QY 477 CAACTTGCAGACGCTGCTGGAGAGGTGAAGCGGGCGTGTTCACATGCGCGCACTTTAT 536
Db 889 CAACTTGCAGACGCTGCTGGAGAGGTGAAGCGGGCGTGTTCACATGCGCGCACTTTAT 948
QY 537 CCGCGCCGACTGCCAGAGTCTGCTACGGGGCATGATCGAGGTGGACGCGCGACGCGGCT 596
Db 949 CCGCGCCGACTGCCAGAGTCTGCTACGGGGCATGATCGAGGTGGACGCGCGACGCGGCT 1008
QY 597 CACGCTAGAGCAATTCAAGAACACATATGTTATAGGGGGCAAGATGAGCCCGAAC 656
Db 1009 CACGCTAGAGCAATTCAAGAACACATATGTTATAGGGGGCAAGATGAGCCCGAAC 1068
QY 657 AGAGCAGCCCATCTCGGAAGGTGCAGATCGCTGCTGCCAGCTCGAGCTGGAGGACATGA 716
Db 1069 AGAGCAGCCCATCTCGGAAGGTGCAGATCGCTGCTGCCAGCTCGAGCTGGAGGACATGA 1128
QY 717 CCGCGACGCTGCTGGAGACGATGCTCACTGGGCTGCTTCGAGACGCGCAACAGCTGCT 776
Db 1129 CCGCGACGCTGCTGGAGACGATGCTCACTGGGCTGCTTCGAGACGCGCAACAGCTGCT 1188
QY 777 GAGGACCTGCTGCTCCGAGGAGGAACACAGGAGAGATGATTTACTTCTCTCTCTGGA 836
Db 1189 GAGGACCTGCTGCTCCGAGGAGGAACACAGGAGAGATGATTTACTTCTCTCTCTGGA 1248
QY 837 CCGGAAAGAAAGGTACCGAGCGAGGAGATGAGACCTGCGCTGCCAGCTCGAGGACATGA 896
Db 1249 CCGGAAAGAAAGGTACCGAGCGAGGAGATGAGACCTGCGCTGCCAGCTCGAGGACATGA 1308
QY 897 CCTTCCCGGAGCGTGTGGACTCCCGATGCTGAACCGGACGAGCGGCGGCGCAGA 956
Db 1309 CCTTCCCGGAGCGTGTGGACTCCCGATGCTGAACCGGACGAGCGGCGGCGCAGA 1368
QY 957 ACGCAATTCATGAGGTGCTCAGCGTGAACGGAACGCGGCTCCCGGTGCTCGCGCGCG 1016
Db 1369 ACGCAATTCATGAGGTGCTCAGCGTGAACGGAACGCGGCTCCCGGTGCTCGCGCGCG 1428
QY 1017 GGCCATTCAGATGGCCGAGCGGACAGGTCTCGGTCCATCAGCGGTGCTCTCAGG 1076
Db 1429 GGCCATTCAGATGGCCGAGCGGACAGGTCTCGGTCCATCAGCGGTGCTCTCAGG 1488
QY 1077 CTTTTCACACGACCTCAGCAGCGCCCGGCTGACCCCTCACCCCTCACCAAGGGCGAG 1136
Db 1489 CTTTTCACACGACCTCAGCAGCGCCCGGCTGACCCCTCACCCCTCACCAAGGGCGAG 1548
QY 1137 TCCCTCTCCCAACCCCAAGGGACACCTGTCCACGCGCAAGGAGAGCCCGGTGGCAC 1196
Db 1549 TCCCTCTCCCAACCCCAAGGGACACCTGTCCACGCGCAAGGAGAGCCCGGTGGCAC 1608
QY 1197 GCCCAACCCCAAGGGCGTCCAGCGCCAGCGTGGAGGGTGCCTTGGAGGGCGCGCT 1256
Db 1609 GCCCAACCCCAAGGGCGTCCAGCGCCAGCGTGGAGGGTGCCTTGGAGGGCGCGCT 1668
QY 1257 CAACTCCATCAAGAACAGCTTCTGGGCTCACCCGCTTCCACCGCGGAACTGCAAGT 1316
Db 1669 CAACTCCATCAAGAACAGCTTCTGGGCTCACCCGCTTCCACCGCGGAACTGCAAGT 1728
QY 1317 TCCGACCGCGGAGAGATGTCAACCTGACACAGAGTGTGCTCCCGAGCTGGCGAAGAA 1376
Db 1729 TCCGACCGCGGAGAGATGTCAACCTGACACAGAGTGTGCTCCCGAGCTGGCGAAGAA 1788
QY 1377 GTCTGGTTTGGGAATCTTCATCAGCCTGGAGAGGAGGAGAGATCTTCTGGTTCATCAA 1436

Db 1789 GTCTGGTTTGGGAATCTTCATCAGCCTGGAGAGGAGGAGAGATCTTCTGGTTCATCAA 1848
QY 1437 AGACAAACCTCTGAGCTCCATCAAGGCTGATCGTGCACGCGCTTCTGTGATTCGAG 1496
Db 1849 AGACAAACCTCTGAGCTCCATCAAGGCTGATCGTGCACGCGCTTCTGTGATTCGAG 1908
QY 1497 TCTCAGCCACAGCGTCACTCTCCAAACGAGCTTCCGGGCGAGTACAAGGCCACGGGGG 1556
Db 1909 TCTCAGCCACAGCGTCACTCTCCAAACGAGCTTCCGGGCGAGTACAAGGCCACGGGGG 1968
QY 1557 GCCAGCCGTTTCAGAAAGCGGTCAGATTCAGGTTGATATCACTACACGAGGGTGG 1616
Db 1969 GCCAGCCGTTTCAGAAAGCGGTCAGATTCAGGTTGATATCACTACACGAGGGTGG 2028
QY 1617 GGAGGCGCAGAGGAGAACGCGATCTACTCCGTCACCTTCCACCTGCTCTCAGGCCCGAG 1676
Db 2029 GGAGGCGCAGAGGAGAACGCGATCTACTCCGTCACCTTCCACCTGCTCTCAGGCCCGAG 2088
QY 1677 CCGTCCGCTTCAAGAGGCTGGTGGAGACCATCCAGGCCAGCTGTGAGCACACACGACCC 1736
Db 2089 CCGTCCGCTTCAAGAGGCTGGTGGAGACCATCCAGGCCAGCTGTGAGCACACACGACCC 2148
QY 1737 GCCTGCGGCCAGCACTTGTGTCAGACACCACTAATCTGATGAAATGATGACGGGCGGCT 1796
Db 2149 GCCTGCGGCCAGCACTTGTGTCAGACACCACTAATCTGATGAAATGATGACGGGCGGCT 2208
QY 1797 TTCCAAATGTGAATTAATCCGAAAAGTTAA 1827
Db 2209 TTCCAAATGTGAATTAATCCGAAAAGTTAA 2239

RESULT 3

US-08-557-006C-38
; Sequence 38, Application US/08557006C
; Patent No. 6258547
; GENERAL INFORMATION:
; APPLICANT: Beri, Rajindar K.
; APPLICANT: Carling, David
; APPLICANT: Forster, Robert A.
; TITLE OF INVENTION: NUCLEIC ACID ENCODING AMP-ACTIVATED PROTEIN KINASE
; FILE REFERENCE: NGAP/PHM37588/UST
; CURRENT APPLICATION NUMBER: US/08/557,006C
; CURRENT FILING DATE: 1996-03-06
; PRIOR APPLICATION NUMBER: PCT/GB94/01093
; PRIOR FILING DATE: 1994-05-20
; PRIOR APPLICATION NUMBER: GB 9310489.1
; PRIOR FILING DATE: 1993-05-21
; PRIOR APPLICATION NUMBER: GB 9318010.7
; PRIOR FILING DATE: 1993-08-31
; NUMBER OF SEQ ID NOS: 44
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 38
; LENGTH: 1742
; TYPE: DNA
; ORGANISM: Human AMP protein kinase
US-08-557-006C-38

Query Match 10.9%; Score 198.8; DB 3; Length 1742;
Best Local Similarity 57.1%; Pred. No. 3.4e-36;
Matches 362; Conservative 0; Mismatches 272; Indels 0; Gaps 0;
QY 4 AAGTGTGAGCGGAGATCGCGATCTTGAAGCTCATTTGAGCACCCCGACGCTTAAAGCTG 63
Db 178 AAAATAAACAGGAATTCAAATCTTAACTTTTCGTCTCTCATATATCAACTC 237
QY 64 CAGACGCTTATGAAACAAAAATATTTGTACTGCTGTAGAACACGTTGTCAGGTGGT 123
Db 238 TACCAAGTGTAGCACTCCACAGACTTTTTTATGTTAATGAAATATGTGTCTGAGGT 297
QY 124 GAGCTCTTCAGCTACCTGTTGAGAGAGGGAGGCTGACGCTTAAGGAGGCTCGAAGTTC 183
Db 298 GAATGTTCAGCTACATCTCTGTAACACACGGAGGGTTGAAGAGGTGGAAGCTCCCGGCTC 357

QY 4 AAGGTGAGCGGAGATCGGATCCTGAAGCTCATTTGAGCACCCCGACGTCCTAAAGCTG 63
Db 201 AAAATAAAGAGAAATTCAAATCTTAACTCTTTCGTCTCATATATCAAACTC 260
QY 64 CAGGAGTTTATGAAACAAAAATATTTGTACCTGGTCTAGAACAGTGTGAGTGT 123
Db 261 TACCAAGTGATCAGCACTCCACAGACTTTTTATGGTAATGGAATATGTGTCTGAGGT 320
QY 124 GAGCTCTTCGACTACTGCTGTGAGAGGGGAGCTGACGCTTAAGGAGGCTGGAGTTC 183
Db 321 GAATTTGTTGCACTACATCTGTAAACACGGGAGGTTGAAGAGGTGGAAGCTCGCGGCTC 380
QY 184 TTCCGCGAGATCATCTCTCGCTGAGACTTCTCCACAGCACTCCATATGCCACAGGAT 243
Db 381 TTCCAGCAGATCTGTCTCGCTGAGACTTCTCAGAGGCACATGTTGTTCACAGGAC 440
QY 244 CTGAACCTGAAACCTCTCTGCTGAGCAGAGAGAACAAATCCGATCGCAGACTTTGGC 303
Db 441 CTGAAGCCAGAGAACGTTGTTGCTGAGCGCCAGATGCTTAAGATAGCTGACTTCGGA 500
QY 304 ATGGGCTCCCTCAGGTTGGGACAGCCTGTTGGAGACAGCTGTGGTCCCGCCACTAC 363
Db 501 CTCTCTAATATGATGTCAGTGTGAATTTCTACGAACTAGTGTGGATCGCCAAATAT 560
QY 364 GCCTGCCGAGGTGATCCGGGGGAGAGATATGACGGCCGGAAGCGGACGCTGTGGAGC 423
Db 561 GCAGCACCGAGGTCTCTCAGGAGGCTGTATGCGGCTCTGAGGTGATATCTGGAGC 620
QY 424 TGGGGGTCTATCTGTTGCCCTTGTGTTGGGGTCTGCCCTTCGACGATGACAACTTG 483
Db 621 TGTGGTGTATCTGTTATGCCCTTCTCTGTGGCACCCTCCGTTGACGATGACGACGTG 680
QY 484 CGACAGCTGCTGAGAGGTGAAGCGGGCGGTGTTCCACATGCCGACTTTATCCGCCC 543
Db 681 CTACGCTCTTAAAGAGATCCGAGGGGTGTCTTACATCCCGAGTATCTCAACCGT 740
QY 544 GACTGCCAGAGTCTGTACGGGCGATGATCGAGGTGGAGCGCGCACCGCCCTCACGCTA 603
Db 741 TCTATTGCCACTCTGCTGATGCACATGCTGCAGGTGGAGCCCTTGAAGCGAGCACTATC 800
QY 604 GAGCATTTCAGAAACATATGTTATATAGGG 637
Db 801 AAAGACATACGAGCATGAATGTTTAAACAGG 834

RESULT 6

US-09-101-146-44
; Sequence 44, Application US/09101146
; Patent No. 6124125
; GENERAL INFORMATION:
; APPLICANT: Dartmouth College, St. Vincent's Institute of
; APPLICANT: Medical Research, Kemp et al.
; TITLE OF INVENTION: No. 6124125el AMP Activated Protein Kinase
; NUMBER OF SEQUENCES: 64
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Jane Massey Licata, Esq.
; STREET: 66 B. Main Street
; CITY: Marlton
; STATE: NJ
; COUNTRY: USA
; ZIP: 08053
; COMPUTER READABLE FORM:
; MEDIUM TYPE: DISKETTE, 3.5 INCH, 1.44 Mb STORAGE
; COMPUTER: IBM PC
; OPERATING SYSTEM: WINDOWS 95
; SOFTWARE: WORDPERFECT 6.0 FOR WINDOWS
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/101,146
; FILING DATE: October 7, 1998
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PN7450

; FILING DATE: 8 JAN 1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Jane Massey Licata
; REGISTRATION NUMBER: 32,257
; REFERENCE/DOCKET NUMBER: DC-0050
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (856) 810-1515
; TELEFAX: (856) 810-1454
; INFORMATION FOR SEQ ID NO: 44:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1647
; TYPE: Nucleic acid
; STRANDEDNESS: Single
; TOPOLOGY: Linear
; ANTI-SENSE: No
; US-09-101-146-44

Query Match 9.9%; Score 181.2; DB 3; Length 1647;
Best Local Similarity 55.8%; Pred. No. 3.3e-32;
Matches 345; Conservative 0; Mismatches 273; Indels 0; Gaps 0;
QY 14 GGGAGATCGGATCCTGAAGCTCATTTGAGCACCCCGACGTCCTAAAGCTGACAGCGTTT 73
Db 188 GAGAGATCCAGAACCTGAAGCTTTTCAGGCAACCTCATATAATCAAACTGTACCAGGTCA 247
QY 74 ATGAAAAACAAAAATATTTGTACCTGCTGTAGAACACGCTGTAGGTGCTGAGCTCTTCG 133
Db 248 TCAGTACACCTCTGATATTTTCATGTTGATGGAATATGTTCTCAGGAGGAGAGCTATT 307
QY 134 ACTACCTGTTGAAGAAGGGGAGGCTGACGCTTAAGGAGGCTCGAAGTTCTTCGCGCAGA 193
Db 308 ATTATATCTGTAAAAATGGAAGTTTGGACGAAAAGGAGAGTGGAGTCTGTTCCAGCAGA 367
QY 194 TCATCTCTGCGCTGGACTTCTGCCACAGCACTCCATATGCCACAGGATCTGAAACCTG 253
Db 368 TCCCTTCTGTTGTTGACTATTTGTCAAGGCATATGTTGGTGTCCACAGAGATTGAAACCTG 427
QY 254 AAAACCTCTCTCGACGAGAGAACAAATCCGCTCGCAGACTTTTGGCATGGGTCCTCC 313
Db 428 AAAAGTCTCTGTTGATGCACACATGATGCAAGATAGCCGACTTCGGTCTTTCAAACA 487
QY 314 TGCAGGTTGGCGACAGCCTGTTGGAGACCACTGTGGGTCCCCCACTACGCTGCCCCG 373
Db 488 TGATGTCAGATGTTGAATTTTAAAGACGAGCTGTGGCTCGCCCAATTTATGCTGCACCA 547
QY 374 AGGTATCCGGGGGAGAGATATGACGGCCGGAAGGGAGCTGTGGAGCTGCGGCTCA 433
Db 548 AAGTAAATTCAGGAAGATTTCTACGCGGCTGAGTAGACATCTGGAGCAGCGGGTCA 607
QY 434 TCCTGTTGCGCTTGTGTTGGGGCTCTGCCCTTCGACGATGACAACTTTCGACAGCTGC 493
Db 608 TTCTATGCTTTGCTGTGTGGAATCTTCCCTTTTGTATGATGACACAGTGCACACTCTTT 667
QY 494 TGGAGAGGTGAAGCGGGGCTGTTCACATGCGCACTTTATCCCGCCGAGTGCACAGA 553
Db 668 TTAAGAAGATATGTGACGGGATTTTTATACCCCTCAGTATTGAAATCCCTCTGTAAATA 727
QY 554 GTCTGCTACGGGGCATGATCGAGTTGGAGCGGCGGACCGGCTCAGCGTAGAGCACTTC 613
Db 728 GCCTTTTGAAGCATATGCTGCAGGTAGTATCTTGAAGAGGGGCCACAAATAAAGATATCA 787
QY 614 AGAAACACATATGTTATA 631
Db 788 GGAACATGAATGTTTA 805

RESULT 7

US-09-984-890-1
; Sequence 1, Application US/09984890
; Patent No. 6492156
; GENERAL INFORMATION:
; APPLICANT: YAN, Chunhua et al.
; TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC

;; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES
;; FILE OF INVENTION: THEREOF
;; FILE REFERENCE: CL001306

;; CURRENT APPLICATION NUMBER: US/09/984,890

;; CURRENT FILING DATE: 2001-10-31

;; NUMBER OF SEQ ID NOS: 4

;; SOFTWARE: FastSeq for Windows Version 4.0

;; SEQ ID NO 1

;; LENGTH: 2175

;; TYPE: DNA

;; ORGANISM: Homo sapiens

US-09-984-890-1

Query Match 9.7%; Score 177.4; DB 4; Length 2175;
Best Local Similarity 54.3%; Pred. No. 2.6e-31;
Matches 358; Conservative 0; Mismatches 301; Indels 0; Gaps 0;

QY 3 GAAGTGGAGCGGAGATCCGATCCTGAAGCTCATTTGACAGCCCGACGCTCTAAAGCT 62
DB 285 GAAACTATTCCGCGAGTAAAGATTAATGAAGGTTTGAATCATCCCAACATAGTTAAATT 344
QY 63 GCACGACGTTTATGAAACAAAAATATTTGTACCTGGTCTAGAACACGCTGTCAGGTGG 122
DB 345 ATTGAAGTGAATTGAGACTGAGAAAAACGCTTACTCTTGTCTATGAGTACGCTAGTGGCGG 404
QY 123 TGAGCTCTTCGACTACCTGCTGAAGAGGGAGGCTGACGCCCTAAGGAGGCTCGGAAGTT 182
DB 405 AGAGGTATTTGATTTACCTAGTGGCTCATGCGAGGATGAAAGAAAGAGGCTCGAGCCAA 464
QY 183 CTTCCGCGAGATCATCTCGGCTGGATCTCTGCCAGAGCATCTCCATATGCCACAGGGA 242
DB 465 ATTCCGCGAGTGTGCTGCTGTCAGTACTCTGTCACAGAAAGTTTATTGTCATAGAGA 524
QY 243 TCTGAACTCTGAAACCTCTGCTGAGAGGAGAGAAACAAACATCCGATCCGACACTTTGG 302
DB 525 CTTAAGGCGAGAAACCTGCTCTTGGATGCTGATATGAACATCAAGATTGCGACACTTTGG 584
QY 303 CATGGCTCCCTGCGAGGTTGGCGACAGCCTGTTGGAGACAGCTGTGGGTCCTCCCACTA 362
DB 585 CTTGAGCAATGAATTCACCTTTGGGAACAGCTGACACCTTCTGTGGCAGTCCCCCTTA 644
QY 363 CGCTGCCCCGAGGTGATCCGGGGGAGAGATGAGCGCGGAGAGCGGACGCTGTGGAG 422
DB 645 TGCTGCCCGACAACTTTCCAGGGCAAAAATATGATGAGACCCGAGGTGATGTGGAG 704
QY 423 CTGCGGGTCACTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 482
DB 705 CCTAGGAGTTATCTCTATACACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 764
QY 483 GCAGACGCTGCTGAGAGGTTGAAGCGGGGCTGCTTCCACATGCCGCACTTTATCCCGCC 542
DB 765 CAAGGAGCTGCGGAAACGCTGCTGAGGGGAAATACCGTATTCATCTACATGTCCAC 824
QY 543 CGACTGCCAGAGTCTGCTACCGGGGATGATCGAGGTGAGCCCGACGCCCTCAGCT 602
DB 825 GGACTGTGAAACCTGCTTAAAGAAATTTCTCATTTTAATCCCAAGAGAGAGCACTTT 884
QY 603 AGACACATTCAGAAACATATGATATAGGGGCAAGATAGGCCGCAACAGAGC 661
DB 885 AGAGCAATCATGAAGATCATGATGATGATGATGATGATGATGATGATGATGATGATGATG 943

RESULT 8

US-08-557-006C-44

;; Sequence 44, Application US/08557006C

;; Patent No. 625847

;; GENERAL INFORMATION:

;; APPLICANT: Beri, Rajindar K.

;; APPLICANT: Carling, David

;; APPLICANT: Forster, Robert A.

;; TITLE OF INVENTION: NUCLEIC ACID ENCODING AMP-ACTIVATED PROTEIN KINASE

;; FILE REFERENCE: NGAP/PHM37588/UST

;; CURRENT APPLICATION NUMBER: US/08/557,006C

;; CURRENT FILING DATE: 1996-03-06

;; PRIOR APPLICATION NUMBER: PCT/GB94/01093

;; PRIOR FILING DATE: 1994-05-20

;; PRIOR APPLICATION NUMBER: GB 9310489.1

;; PRIOR FILING DATE: 1993-05-21

;; PRIOR APPLICATION NUMBER: GB 9318010.7

;; PRIOR FILING DATE: 1993-08-31

;; NUMBER OF SEQ ID NOS: 44

;; SOFTWARE: PatentIn Ver. 2.1

;; SEQ ID NO 44

;; LENGTH: 1747

;; TYPE: DNA

;; ORGANISM: Human

;; FEATURE:

;; NAME/KEY: gene

;; LOCATION: (1)-(1747)

;; OTHER INFORMATION: Full length cDNA sequence fragment of Human AMPK -

;; OTHER INFORMATION: fragment begins at nucleotide 24 and ends with

;; OTHER INFORMATION: nucleotide 1765

US-08-557-006C-44

Query Match 9.5%; Score 173.8; DB 3; Length 1747;

Best Local Similarity 56.7%; Pred. No. 1.6e-30; Indels 5; Gaps 2;

Matches 362; Conservative 0; Mismatches 272;

QY 4 AAGGTGGAGCGGAGATCCGATCCTGAAGCTCATTTGAGCACCCCGACGCTCTAAAGCTG 63
DB 178 AATAAATAACAGAGAAATTCAAAATCTTAAACTCTTTCGTATCTCTATATTAACAATC 237
QY 64 CACGAGCTTTATGAAAAAATAAATAATTTGTACCTGGTCTAGAACACGCTGTCAGGTGT 123
DB 238 TACCAAGTGATCAGCACTTCAACAGACTTTTTTATGGTAATGGAATATATGTCTGAGGT 297
QY 124 GAGCTCTTCGACTACTGCTGTAAGAGGGAGGCTGACGCCCTAAGAGGCTCGGAAGTTC 183
DB 298 GAATTGTCGACTACATCTGTAAAACAGGAGGGTTGAAGAGTGAAGCTCGCGGCTC 357
QY 184 TTCGGCAGATCATCTCTGCGCTGGACTTCTG - CCACAGCCACTCCATATGCCACAGGG 241
DB 358 TTCCAGCAGATTCTGTCTGCGTGGACTACTGTCTCACAGGCACATGTTGTCCACAGGG 417
QY 242 ATCTGAAACCTGAAACCTCTCTGCTGGAGAGAGAGAACACATCCGCATCGCAGACTTGG 301
DB 418 ACCTGAAGCCAGAGAACGTTGCTGGACGCCAGATGAATGCTAAGATAGTGAATCTCG 477
QY 302 GCATGCGCTCCCTGCGAGGTGGCGACAGCTGTTGGAGACACAGCTGTGGTCTCCGCCACT 361
DB 478 GACTCTCTAATATGATGTCAGATGTTGAATTTCTACGAATAGTGTGGATCGCAAAT 537
QY 362 ACCTCTGCCCCGAGGTGATCCGGGGGGAGAGTATGACCGCGGAAAGGCGGAGCTGTGGA 421
DB 538 ATGCAGCACCGGAGGTCACTCAGGAAGGCTGTATCGGGTCTCTGAGGTGATATCTGGA 597
QY 422 GCTGCGGCTCATCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 478
DB 598 GCTGTGTTGTTATCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 657
QY 479 ACTTGGACAGCTGCTGGAGAGGTGAAAGCGGGCGTGTTCACATGCGGCACTTTATCC 538
DB 658 ACCTGCTACGCTCTTTAAGAGATCCGAGGGGGTGTGTTCTACATCCCGGAGTATCTCA 717
QY 539 CGCCCCGACTGCCAGATCTGCTACGGGGGATGATCGAGGTGAGCGCGCACCGCGCTCA 598
DB 718 ACCGTCTATTGGCACTCTGCTGATGCACATGCTGCAGGTGGACCCCTTTGAGCGAGCA 777
QY 599 CGCTAGACCATTCAGAAACACATATGTTATAGGG 637
DB 778 CTATCAAGACATACGAGAGCATGAATGGTTTAAACAGG 816

RESULT 9

US-09-359-161-4

;; Sequence 4, Application US/09359161A

Patent No. 6342656
; GENERAL INFORMATION:
; APPLICANT: Bradford, Kent J.
; APPLICANT: Dahal, Peetambar
; APPLICANT: Yang, Hong
; APPLICANT: Cooley, Michael
; APPLICANT: Downie, Bruce
; APPLICANT: Gee, Oliver
; APPLICANT: The Regents of the University of California
; TITLE OF INVENTION: Regulation of Source-Sink Relationships and Responses
; FILE REFERENCE: 021070-095900US
; CURRENT APPLICATION NUMBER: US/09/359,161A
; CURRENT FILING DATE: 1999-07-21
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 4
; LENGTH: 1929
; TYPE: DNA
; ORGANISM: Lycopersicon esculentum
; FEATURE:
; OTHER INFORMATION: Lycopersicon esculentum plant homolog of yeast
; OTHER INFORMATION: SNF1 kinase subunit of protein kinase (LeSNP1)
US-09-359-161-4

Query Match 9.4%; Score 171.6; DB 4; Length 1929;
Best Local Similarity 54.6%; Pred. No. 5.3e-30;
Matches 342; Conservative 0; Mismatches 284; Indels 0; Gaps 0;
QY 3 GAAGTGGAGCGGAGATCGGATCCTGAAGCTCAATTGAGCACCCACGCTCTAAAGCT 62
DB 218 GAAAGTCGTAGAGAAATCAAAATATTGAGATTGTTCATGCTCATATATATACGGCT 277
QY 63 GCACGACGTTTATGAAAAAATAATTTGTACCTGTGTGTAGAACACGCTGTCAGGTGG 122
DB 278 TTATGAGTCTAGAGACACCATCAGATATATATGTTGTGATGAGTATGAAATCTGG 337
QY 123 TGAGCTTTTCAGTACTCTGCGTGAAGAGGAGGCTGACGCTTAAGGAGGCTCGGAAGTT 182
DB 338 CGAGTTATTTGATTACATTTGAGAGGCGAGATTGCGAGGAGTGAAGCTCGTAACCT 397
QY 183 CTTCCGCGAGATCATCTGCGTGGACTTTGCCAGACCACTCCATATGCCACAGGA 242
DB 398 TTTTCAGCAGATAATTTCTGTGTGGAGTACTGCCATAGAAACATGCGTTCATAGAGA 457
QY 243 TCTGAAACCTGAAACCTCCTGCTGGACGAGAGAAACAACTCCGATCGCAGACTTTGG 302
DB 458 CTTAAGCTGAAACCTCCTCTGGAATCCAAATGGAATGTGAAGTTCGAGATTTGG 517
QY 303 CATGCGTCCCTGCAAGTTGGCGACAGCCTGTTGGAGACCACTGTTGGTCCCCCACTA 362
DB 518 TTTGAGCAATATCATGCGCGATGGTCAATTTCTGAAGACAAAGTTGCGGAAGCCCAACTA 577
QY 363 CGCTGCCCCGAGGTGATCCGGGGGAGAGTATGACCGCGGAGGCGGAGCTGTGGAG 422
DB 578 TGCTGCCCCAGAGTTATATCAGGTAATTTGATGCTGGCCCTGAGTAGATGATGGAG 637
QY 423 CTGCGGCTCATCTGTTGCTGCTGCTGGGGGCTCTGCGCTTCGACGATGACAACTT 482
DB 638 CTGTGTTGTTATCTTTATGTTCTCTCTGTGGACACCTTCGGTTGAGATGAAACAT 697
QY 483 GCGACAGTCTGTGGAGAGAGGTGAAGCGGGCGGTGTTCCACATGCGCGCACTTTATCCGCC 542
DB 698 ACCCAATCTTTTAAGAAAAATAAGGGTGAATATATACTCTGCGCCAGCCATTTATCAGC 757
QY 543 CGACTGCCAGAGTCTGCTACGGGGATGATCGAGTGGACCGCGCACCGCCCTCAGCT 602
DB 758 TGGTGGAGGGATTTGATTCGAGGATGCTTATAGTCGACCAATGAAGCGAATGACTAT 817
QY 603 AGAGCACATTCAGAAACACATATGCT 628
DB 818 TCCTGAGATTGCGCTGACCCCTTGGT 843

RESULT 10
US-09-579-664B-4
; Sequence 4, Application US/09579664B
; Patent No. 6514719
; GENERAL INFORMATION:
; APPLICANT: Immunex Corporation
; APPLICANT: Bird, Timothy A.
; APPLICANT: Virca, G. Duke
; APPLICANT: Martin, Unja
; APPLICANT: Anderson, Dirk M.
; TITLE OF INVENTION: NOVEL MURINE AND HUMAN KINASES
; FILE REFERENCE: 2923-A
; CURRENT APPLICATION NUMBER: US/09/579,664B
; CURRENT FILING DATE: 2000-05-26
; NUMBER OF SEQ ID NOS: 36
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 4
; LENGTH: 2902
; TYPE: DNA
; ORGANISM: Mus musculus
US-09-579-664B-4

Query Match 9.2%; Score 168.4; DB 4; Length 2902;
Best Local Similarity 55.0%; Pred. No. 3.2e-29;
Matches 353; Conservative 0; Mismatches 286; Indels 3; Gaps 1;
QY 2 TGAAGTGGAGCGGAGATCGGATCCTGAAGCTCATTGAGCACCCACGCTCTAAAGC 61
DB 418 TGCATCATACGAGGAGATTGAGATCATGCTTCTCACTCAACCCCCACATCATTTGCCA 477
QY 62 TGACAGCAGTTTATGAAAAAATAATTTGTACCTGTGTGTAGAACACAGCTGTCAAGTG 121
DB 478 TCCATGAAGTTTGAATAGCAGCAAGATTGTGATTGTTCATGAGTATGCCAGCCGAG 537
QY 122 GTGAGCTTTCGACTACTCTGCTGAAAGAGGAGGCTGACGCTTAAGAGGCTCGGAAGT 181
DB 538 GCGATCTGTATGATTATCATCAGTGAGCGGCCACGCGCTGAGTGAGCGGAGCGCCAGGCA 597
QY 182 TCTTCGCGGAGATCATCTGCGCTGGACTTTGCCACAGCACTCCATATGCCACAGGG 241
DB 598 TCTTCGAGAGATGCTGTCTGCTGCACTACTGCCACAGAACGAGATCGTTCAACGAG 657
QY 242 ATCTGAAACCTGAAACCTCCTGCTGGACGAGAGAAACAACTCCGATCGCAGACTTTG 301
DB 658 ATCTCAAGCTGGAACACATCCTTCTAGATGCAATGGAACACATCAAGATTGCTGACTTG 717
QY 302 GCATGGCGTCCCTGAGGTTGGCGACAGCCTGTTGGAGACAGCTGTGGTCCCCCACT 361
DB 718 GCCTCTCAACCTGTATACCAAGGCAAGTTCTCTCCAGAGCTTCTGTGGGAGCGCTCTCT 777
QY 362 ACGCTGCCCGAGGTGATCCGGGGGAGAAATATGACGCGCGGAGCGGACGCTGTGGA 421
DB 778 ACGCTCGCTGAGATAGTCAACGGGAAGCCCTATGTGGGCCCCAGAGGTGGAAGCTGGT 837
QY 422 GCTGCGCGCTCATCTGCTGCTGCTGCTGGGGGCTCTGCGCTTCGACGATGACAACT 481
DB 838 CTCTGGCGGTTCTCTCTGTACATCTGCTGATGSCACCATGCCCCCTTTGACGGGAGGATC 897
QY 482 TGCGACAGTCTGTGGAGAGGTGAGCGGGGCGGTGTTCCATGCTCCGCACTTTATCCGCG 541
DB 898 ATAAACACCTGGTGAAGCAATCAGTAACGGGGGCTTACCGTGAAGCGGCCCC---AAGCGT 954
QY 542 CCGACTGCCAGAGTCTGCTACGGGGCATGATCGAGGTGGAGCGCGCGCACCGCTCAGCG 601
DB 955 CCGATGCTGTGGCTGATCGGTGGCTGTTAATGTTGAACCCACCGCTCGGCCACAC 1014
QY 602 TAGAGCACATTCAGAAAACACATATGTTATAGGGGGCAAGA 643
DB 1015 TGGAGGATGTAGCCAGTCAATTGGTGGTCAACTGGGGTTTACA 1056

RESULT 11

```

US-08-557-006C-37
; Sequence 37, Application US/08557006C
; Patent No. 6258547
; GENERAL INFORMATION:
; APPLICANT: Beri, Rajindar K.
; APPLICANT: Carling, David
; APPLICANT: Forder, Robert A.
; TITLE OF INVENTION: NUCLEIC ACID EN
; FILE REFERENCE: NGAP/PHM37588/JUST
; CURRENT APPLICATION NUMBER: US/08/555
; CURRENT FILING DATE: 1996-03-06
; PRIOR APPLICATION NUMBER: PCT/GB94/
; PRIOR FILING DATE: 1994-05-20
; PRIOR APPLICATION NUMBER: GB 931048
; PRIOR FILING DATE: 1993-05-21
; PRIOR APPLICATION NUMBER: GB 931801
; PRIOR FILING DATE: 1993-08-31
; NUMBER OF SEQ ID NOS: 44
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 37
; LENGTH: 1736
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of A
; OTHER INFORMATION: protein kinase
US-08-557-006C-37

```

Query Match	8.3%	Score	151.6	DB	3	Length	1736
Best Local Similarity	52.5%	Pred. No.	1.8e-25				
Matches	331	Conservative	0	Mismatches	299	Indels	0
Gaps	0						
Qy	4	AAGGTGAGCGGGAGATCGCGATCCTGAAGCTCATTTGAGCACGCCACCCACGCTCTAAAGCTG	63				
Db	178	AAATAAACAAGAGAAATCAAAATCTAAACTCTTTGGTCATCTCATATATCAAACTA	237				
Qy	64	CAGCAGTTTATGAAAAACAAAAATATTTGTACTGTGTGTAGAAACAGTGTCTAGGTGGT	123				
Db	238	TACCAGGTGATCAGCACTCCCAACAGATTTTTTTTATGGTAATGGAATATGTCTCTGAGGT	297				
Qy	124	GAGCTCTTCGACTACCTTGGTGAAGAAGGGAGGCTGACGCCCTAAGGAGGCTCGGAAGTTC	183				
Db	298	GAATTTTGTACTACATCTGTAAAGATGGACGGTTGAAGAGATGGAAGCCAGCGGCTC	357				
Qy	184	TTCCGGCAGATCATCTCTGCGCTGGACTTTTGCACAGCCACTCCATATGCCACAGGGAT	243				
Db	358	TTTCAGCAGATTCTGTCTGTCTGGATTACTGTCTATAGGCATATGGTTGTTTCATCGAGAC	417				
Qy	244	CTGAAACCTGAAACCTCTCTCTGGACGAGAGAACAACATCCGCATCGCAGACTTTGGC	303				
Db	418	CTGAAACGAGAGAATGCTCTTTGGATGCAACATGAATGCAAGATAGCCGATTCGGGA	477				
Qy	304	ATGCGCTCCCTGCAGGTGGCGACAGCCCTGTGTGGAGACCAGCTGTGGGTCCCCCACTAC	363				
Db	478	TTATCTAATATGATGTGAGATGGTGAATTTCTGAGAACTAGTTGCGGATCTCCAATATT	537				
Qy	364	GCCTGCCCGAGGTGATCCGGGGGGAAGTATGACGGCGGAAGGCGAGCTGTGAGC	423				
Db	538	GCAGCACTGAAGTCACTCTCAGGCAGATTTGTATCGAGTCTGAAGTGTGATATCTGGAGC	597				
Qy	424	TGCGGCGTCATCTGTTTCGCTGTGTGTGGGGCTCTGCCCTTTCGACGATGACAACTTG	483				
Db	598	TGTGGTGTATCTTGATATGCTCTCTTTGTGGCACCCCTCCCAATTTGATGAGCATGTA	657				
Qy	484	CGACAGCTGTGGAGAAGGTGAAGCGGGGCGTGTTCACATGCGCGCACTTTATCCCGGCC	543				
Db	658	CCTACGTTATTTAAGAAGATCCGAGGGGGTCTCTTTTATATCCCAGAATATCTCAATCGT	717				
Qy	544	GACTGCCAGACTCTGCTACGGGGCATGATCCAGGTGGAACGGCGCACCGCCTCACGCTA	603				
Db	718	TCTGTGCGCACTCTCCTGATGATATGCTGCAGGTTGACCCACTGAAACGAGCAACTATC	777				
Qy	604	GAGCACATTTGAGAAACACATATGGTATATA	633				

Db 778 AAGACATAGAGCATGAATGGTTAAA 807

RESULT 12

US-08-557-006C-36

; Sequence 36, Application US/08557006C

; Patent No. 6258547

; GENERAL INFORMATION:

; APPLICANT: Beri, Rajindar K.

; APPLICANT: Carlving, David

; APPLICANT: Forlder, Robert A.

; TITLE OF INVENTION: NUCLEIC ACID ENCODING AM

; FILE REFERENCE: NGAP/PRM37588/UST

; CURRENT APPLICATION NUMBER: US/08/557,006C

; CURRENT FILING DATE: 1996-03-06

; PRIOR APPLICATION NUMBER: PCT/GB94/01093

; PRIOR FILING DATE: 1994-05-20

; PRIOR APPLICATION NUMBER: GB 9310489.1

; PRIOR FILING DATE: 1993-05-21

; PRIOR APPLICATION NUMBER: GB 9318010.7

; PRIOR FILING DATE: 1993-08-31

; NUMBER OF SEQ ID NOS: 44

; SOFTWARE: Patentin Ver. 2.1

; SEQ ID NO 36

; LENGTH: 1783

; TYPE: DNA

; ORGANISM: Human AMP protein kinase

US-08-557-006C-36

Query Match	8.3%;	Score 151.6;	DB 3;	Length 1783;
Best Local Similarity	52.5%;	Pred. No. 1.8e-25;		
Matches 331;	Conservative 0;	Mismatches 299;	Indels 0;	Gaps 0;
QY	4	AAGGTGAGCGGGAGATCGCGATCCTCAAGCTCATTTGACGACACCCACGTCCTCAAGCTG	63	
DB	179	AAATATAACGAGAAATCAAAATCTAAACTCTTTTGGTCATCTCATATTAACAATA	238	
QY	64	CACGACGTTTATGAAAAACAAAAATATTTGPACTGTGCTGCTAGAACACGCTGTCAAGTGGT	123	
DB	239	TACCAGGTGATCAGCACTCCAAACAGATTTTTTATGTGTAATGGAATATGTCCTGGAGGT	298	
QY	124	GAGCTCTTCGACTACCTGGTGAACAGGGGAGGCTGACGCCCTAAGGAGGCTCGGAAGTTC	183	
DB	299	GAATATTTGACTACATCTGTAAGCATGAGCGGTTGAAGAGATGGAAGCAGCGGGCTC	358	
QY	184	TTCCGGCAGATCATCTCTGCGCTGGACTTCTGCCACAGCCACTCCATATGCCACAGGGAT	243	
DB	359	TTTCAGCAGATTCTGCTGCTGTGGATTACTGTCTAGGCATATGGTTGTTTCATCGAGAC	418	
QY	244	CTGAAACCTGAAACCTCTCTCTGGACGAGAGAAACAACATCCCGATCGCAGACTTTGGC	303	
DB	419	CTGAAACGAGAGAAATGCTCTGTTGGATGACACATGAATGCAAGATAGCCGATTTCCGA	478	
QY	304	ATGCGCTCCCTGCAGGTTGGCGACAGCCTGTTTGGAGACCACTGTGGGTGCCCCCACTAC	363	
DB	479	TTATCTAATATGATGTGATGTGTAATCTTCGAAACTAGTTCGGGATCTCCAAATAT	538	
QY	364	GCCTGCCCCGAGGTGATCCGGGGGAGAAAGTATACGCGCCGGAAGCGGACGTGTGGAGC	423	
DB	539	GCAGCACCTGAAGTCATCTCAGGCAGATTGTATGCGAGTCTGAAGTTGATATCTGGAGC	598	
QY	424	TGGCGGTCATCCTGTTCGCTTCTGCTGGTGGGGGCTCTGCCCTTCGACGATGACAACTTG	483	
DB	599	TGTGGTGTATCTTTGTATGCTCTCTTTGTGGCACCCCTCCCATTTGATGATGAGCATGTA	658	
QY	484	CGACAGCTGCTGGAAGAAGTGAACGGGGGCGGTGTTCACATGCGGCACTTTTATCCCGGCC	543	
DB	659	CCTAGTTATTATAGAGATCCGAGGGGGTGTCTTTTATATCCCAAGATATCTCAATCGT	718	
QY	544	GACTGCCAGACTCTGTCTACGGGGGCAATGATCGAGGTGGAGCCGCCACGCCGCTCTACGCTA	603	
DB	719	TCTGTGCGCACTCTCTGATGCATATGTGCAAGTTGACCCACTGAAACGAGCACTATC	778	

Db 426 CCTGAAGCCCGAGAAATCATCTGCTGGACAAGCAGCAGCCCGCAATTAAGCT 485
Qy 291 CGCAGACTTTGGCATGGCTCCCTGCGAGTTGGCGACAGCTTGTGGAGACCACTGTGG 350
Db 486 CATCGACTTTGGCATCGCGACAGATCGAGGCTGGCAGGATTCAGAAACATCTTTGG 545
Qy 351 GTCCCCCACTAGCGCTGCCCCGAGGTGATCCGGGGGGAAGTATGACGGCCGGAAGGC 410
Db 546 CACACCCGAGTTTGTCCCCCGA--GATCGTGAACATATGAGCCACTTTGGCTTGGAGGC 602
Qy 411 GGAGCTGTGAGCTGGCGGTCACTCTGCTTGGCTTGGCTTGGCTTGGCTTGGCTTGG 470
Db 603 TGACATGTGAGCATTTGGCGTCACTACCTACATCTCTCTGAGCGGAGCGTCCCATTCCT 662
Qy 471 CGATGACAACTTGGCAGCTGTGGAGAA 500
Db 663 GGGCGAGACCAAGCAGGAGCGCTGACGAA 692

RESULT 15
US-09-186-277-4
; Sequence 4, Application US/09186277
; Patent No. 6171841
; GENERAL INFORMATION:
; APPLICANT: AKIRA, SHIZUO
; APPLICANT: KAWAI, TARO
; TITLE OF INVENTION: DNA CODING FOR SERINE/THREONINE KINASE
; FILE REFERENCE: 081356/0128
; CURRENT APPLICATION NUMBER: US/09/186,277
; EARLIER FILING DATE: 1998-11-05
; EARLIER APPLICATION NUMBER: JP97/261589
; EARLIER FILING DATE: 1997-09-26
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 4
; LENGTH: 1429
; TYPE: DNA
; ORGANISM: Mus musculus
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (10)..(1353)
US-09-186-277-4

Query Match 7.4%; Score 135.2; DB 3; Length 1429;
Best Local Similarity 57.3%; Pred. No. 9.2e-22;
Matches 292; Conservative 0; Mismatches 203; Indels 15; Gaps 2;
Qy 3 GAAGTGGAGCGGAGATCGCGATCCTGAGCTCATTTAGCAGCAGCCCGCACTCTAAAGCT 62
Db 186 GGAGATCGAACCGAGGAGTGCATCTCGCGAGATCGCCCAACATCATACACT 245
Qy 63 GCAGAGCTTTATGAACAAATAATTTCTGCTGTAGAACACGCTCAGGTGG 122
Db 246 GCATGACGTGTTGAGAACAAACAGATGTGTGCTGATCTTGGAGCTGGTTCGGTGG 305
Qy 123 TGAGCTCTTCGACTACCTGGTGAAGAGGAGGCTGACGCTTAAGGAGGCTCGGAAGTT 182
Db 306 CGAGCTTTTCGACTTCTCTGGCGAGAGGAGTTCATTGACGGAGGATGAGGCCACGAGTT 365
Qy 183 CTTCCGGCAGATCATCTCTGCGCTGGACTTTGCGACAGCCACTCCATATGCCACAGGGA 242
Db 366 CCTCAAAACAAATCCTAGACGGGTGTCACCTACCTGCACTCCAAAGCGCATCGCACACTTTGA 425
Qy 243 TCTGAAACCTGAAACCTCTCTGCTGCGAGAGAAACATATCCGC-----AT 290
Db 426 CTTGAAGCCCGAAGAACATATGTTCTGGAACAGCAGCAGCCCGCCCATTAAGCT 485
Qy 291 CGCAGACTTTGGCATGGCTCCCTGCGAGTTGGCGACAGCTTGTGGAGACCACTGTGG 350
Db 486 CATCGACTTTGGCATCGCGCAGGATCGAGGCTGGCAGGAGTTCAAGAACATCTTTGG 545
Qy 351 GTCCCCCACTAGCGCTGCCCCGAGGTGATCCGGGGGGAAGTATGACGGCCGGAAGGC 410

Db 546 CACACCCGAGTTTGTCCCCCGA--GATCGTGAACATATGAGCCCACTTGGCTTGGAGGC 602
Qy 411 GGAGCTGTGAGCTGGCGGTCACTCTGCTTGGCTTGGCTTGGCTTGGCTTGGCTTGG 470
Db 603 TGACATGTGAGCATTTGGCGTCACTACCTACATCTCTCTGAGCGGAGGCTCCCATTCCT 662
Qy 471 CGATGACAACTTGGCAGCTGTGGAGAA 500
Db 663 GGGCGAGACCAAGCAGGAGCGCTGACGAA 692

Search completed: November 27, 2003, 11:09:42
Job time : 121.131 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 27, 2003, 04:08:07 ; Search time 3468.16 Seconds
(without alignments)
12803.430 Million cell updates/sec

Title: US-10-054-579-3

Perfect score: 1827

Sequence: 1 atgaagtgaggcggagat.....gaattatcccgaaagtaa 1827

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 22781392 seqs, 12152238056 residues

Total number of hits satisfying chosen parameters: 45562784

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

```
EST:*
1: em_estba:*
2: em_esthum:*
3: em_estin:*
4: em_estmu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_htc:*
9: gb_est1:*
10: gb_est2:*
11: gb_htc:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estfun:*
16: em_estom:*
17: em_gss_hum:*
18: em_gss_inv:*
19: em_gss_pln:*
20: em_gss_vrt:*
21: em_gss_fun:*
22: em_gss_mam:*
23: em_gss_mus:*
24: em_gss_pro:*
25: em_gss_rod:*
26: em_gss_phg:*
27: em_gss_vrl:*
28: gb_gss1:*
29: gb_gss2:*
```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	639.4	35.0	769	13	BQ442940
2	601.4	32.9	732	13	BQ443783
3	599	32.8	791	13	BU611869
4	595.2	32.6	1201	9	AL538014

5	593.4	32.5	723	14	CA315127	CA315127	UI-M-FW0-
6	588.2	32.2	733	13	BQ178695	BQ178695	UI-M-EV0-
7	584.8	32.0	733	10	BG395625	BG395625	602458251
8	583.8	32.0	1882	11	BC017182	BC017182	Homo sapi
9	574.2	31.4	712	13	BQ180079	BQ180079	UI-M-EW0-
10	562.2	30.8	740	10	BF222739	BF222739	7q32f02.x
11	559.8	30.6	563	10	BE251924	BE251924	601107590
12	548.2	30.0	693	13	BQ443644	BQ443644	UI-M-EW0-
13	515	28.2	589	10	BE259121	BE259121	601107773
14	496.8	27.2	699	14	CA319314	CA319314	UI-M-FW0-
15	495.2	27.1	727	12	BI925755	BI925755	603072158
16	478.8	26.2	746	13	BQ394353	BQ394353	603804548
17	476.8	26.1	697	13	BQ392070	BQ392070	603802721
18	459	25.1	889	13	BQ434571	BQ434571	AGENCOURT
19	445.6	24.4	709	12	BI916891	BI916891	603177863
20	445.2	24.4	546	12	BM944658	BM944658	UI-M-EHOP
21	445	24.4	937	10	BF527021	BF527021	602039908
22	425.8	23.3	947	10	BF525860	BF525860	602070128
23	425.6	23.3	881	13	BQ717903	BQ717903	AGENCOURT
24	421.6	23.1	698	12	BM948524	BM948524	UI-M-EGOP
25	413.8	22.6	1041	13	BQ882776	BQ882776	AGENCOURT
26	411.2	22.5	886	13	BU285204	BU285204	603601981
27	406.2	22.2	701	10	BE256016	BE256016	601113531
28	405.4	22.2	733	9	AI135353	AI135353	GH13047.5
29	403.6	22.1	831	10	BF529743	BF529743	602044002
30	398	21.8	1177	12	BM466629	BM466629	AGENCOURT
31	393.6	21.5	639	10	BF342303	BF342303	602013084
32	392.6	21.5	600	13	BQ444032	BQ444032	UI-M-EW0-
33	388.2	21.2	651	14	CA368080	CA368080	644147 NC
34	387	21.2	476	14	CD355265	CD355265	UI-M-FY0-
35	385.6	21.1	477	13	BQ554978	BQ554978	H4031D07-
36	379	20.7	742	12	BM944329	BM944329	UI-M-EHOP
37	377	20.6	378	13	EX090836	EX090836	EX090836
38	362.8	19.9	728	13	BQ769984	BQ769984	UI-M-PIO-
39	355.4	19.5	898	10	BF529174	BF529174	602042051
40	355.2	19.4	783	13	BU348106	BU348106	603524982
41	352	19.3	1073	12	BM475434	BM475434	AGENCOURT
42	339.8	18.6	730	13	BU613922	BU613922	UI-M-FRO-
43	337.6	18.5	507	10	BF461585	BF461585	UI-M-CGOP
44	329	18.0	743	13	BU362743	BU362743	603786550
45	325.4	17.8	614	12	BM944236	BM944236	UI-M-EHOP

ALIGNMENTS

RESULT 1
BQ442940
LOCUS
DEFINITION
UI-M-EV0-bxf-o-02-0-UI.r1 NIH-BMAP_EV0 Mus musculus CDNA clone
IMAGE:5707609 5', mRNA sequence.
769 bp mRNA linear EST 29-MAY-2002
ACCESSION
BQ442940
VERSION
BQ442940.1 GI:21246052
KEYWORDS
EST.
SOURCE
Mus musculus (house mouse)
ORGANISM
Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE
1 (bases 1 to 769)
AUTHORS
NIH-MGC <http://mgc.nci.nih.gov/>.
TITLE
National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL
Unpublished
COMMENT
Contact: Robert Strausberg, Ph.D.
Email: cgabbs@email.nih.gov
Tissue Procurement: Dr. James Lin, University of Iowa
CDNA Library preparation: Dr. M. Bento Soares, University of Iowa
CDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
This clone was contributed by the Brain Molecular Anatomy Project
(BMAP)


```

Seq primer: pYX-5,
Location/Qualifiers
1. 769
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6"
/db_xref="taxon:10090"
/clone="IMAGE:5707609"
/tissue_type="whole brain"
/dev_stage="embryo 15.5 dpc"
/lab_host="DH10B (TI phage resistant)"
/clone_lib="NIH_BMAP_EVO"
/notes="Organ: brain; Vector: pYX-Asc; Site 1: EcoR I; Site 2: Not I; The library was constructed according to Bonaldo, Lennon and Soares, Genome Research, 6:791-806, 1996. Denatured mRNA was size fractionated on a 1% agarose gel. First strand cDNA synthesis was primed with an oligo-dT primer containing a Not I site. Double stranded cDNA was size selected according to mRNA size fraction, ligated with EcoR I adaptor, digested with Not I, and then cloned directionally into pYX-Asc vector. The library tag sequence located between the Not I site and the polyA tail, is GTGCGTGGAA. This library was created for the University of Iowa Mouse Brain Molecular Anatomy Project (BMAP): 'Gene Discovery in the Developing Mouse Nervous System', supported by National Institutes of Mental Health (NIMH), Hemin Chin, Ph.D., program coordinator."

BASE COUNT 185 a 241 c 202 g 141 t
ORIGIN

Query Match 35.0%; Score 639.4; DB 13; Length 769;
Best Local Similarity 89.5%; Pred. No. 8.8e-131;
Matches 688; Conservative 0; Mismatches 81; Indels 0; Gaps 0;

Qy 721 GACGTGTCGACAGCATGCACTCACTGGCTGCTTCGAGACCCCAACAGACTGCTCCAG 780
Db 1 GATGTGTGGACAGCATGCACTCACTGGCTGCTTCGAGACCCCAACAGACTGCTCCAG 60
Qy 781 GACGTGTCGACAGGAGGAGCAACAGAGAGATGATTCTTCCTCTCTGGACCGG 840
Db 61 GATGTGTCATCTGAGGAGGAGATCAGAAAGATGATTATTTCTCTCTGGATCGG 120
Qy 841 AAAGAAAGGTACCCGAGCAGGAGGATGAGGACCTGCCCTCCCGGAAACGAGATAGACCT 900
Db 121 AAAGAAAGGTATCCAAAGCATGAGATGAGGACCTGCCCTCCCGGAAACGAGATAGACCT 180
Qy 901 CCCGGAAGCTGTGGACTCCCGCATGCTGAACCGGACGCGGACGCGGCGGACGACGC 960
Db 181 CCCGGAAGCTGTGGATTCCTCCGATGCTGAACCGGACGCGGACGCGGACGCTGAGCGC 240
Qy 961 AAATCCATGGAAGTGTCTCAGCGTGACGACGCGGCTCCCGGTGCTTCGCGCGGCGGCC 1020
Db 241 AAGTCCATGGAAGTGTCTCAGTGTACAGATGCTGCTCCCGGTGCTTCGCGGAGAGCC 300
Qy 1021 ATTGAGATGGCCACGACGCGGACGAGTCTCGGTCCATCAGCGTGTGCTCTCAGGCTT 1080
Db 301 ATTGAGATGGCCACGACGCGGACGAGATCTCGATCCATCAGTGTGCTCTCAGGCTT 360
Qy 1081 TCCACCGCCCACTCAGACGCGGCTGACCCCTCACTCCCTCACTCCAGGGGAGTCCC 1140
Db 361 TCTCAAGTCACTCAGACGCTCTCGGTGACCCCTCACTCCCTCACTCCAGGGGAGTCCC 420
Qy 1141 CTCCCGCCACCCAGGGGACACTGTCCACACGCGGAGAGAGAGAGAGAGAGAGAGAGAG 1200
Db 421 CTTCCTACCCCGGAGGAGCGCTGTCCACAGCGGAGAGAGAGAGAGAGAGAGAGAGAG 480
Qy 1201 AACCCACGCGCCCGTCCAGCGGAGCGTGTGAGGGGTGCTTCGAGGGGCGCGCTCAAC 1260
Db 481 AACCCACGCGCCCGTCCAGCGGAGTGTGAGGGGTGCTTCGAGGGGCGCGCTCAAC 540
Qy 1261 TCCATCAGAACAGCTTTCTGGGCTCACTCCCGCTTCACCGCGGAACTCAAGTTCCG 1320
Db 541 TCCATCAGAACAGCTTTCTGGGCTCACTCCCGCTTCACCGCGGAACTCAAGTTCCG 600

```

1321 ACGCCGAGGAGATGTCACACCTGACACAGAGTCGTCCCGAGAGCTGCGGAGAGTCC 1380
 Db 601 ACGCCGAGGAGATGTCACACCTGACACCGAGATCTCTCCAGAGCTGCGGAGAGTCC 660
 Qy 1381 TGGTTTGGGAACTTCATCAGCTCGAGAGGAGGAGGAGAGATCTTCGTGTCATCAAGAC 1440
 Db 661 TGGTTTGGGAACTTCATCAGCTCGAGAGGAGGAGGAGAGATCTTCGTGTCATCAAGAC 720
 Qy 1441 AAACCTCTGAGCTCCATCAGCTGACATCGTCGACGCTTCCTCTCGA 1489
 Db 721 AAACCTCTGAGCTCCATCAGCTGACATCGTCGTCATCGCTCTCTCGA 769

RESULT 2
 BQ443783
 LOCUS
 DEFINITION
 IMAGE:5708522 5', mRNA sequence.
 ACCESSION
 BQ443783
 VERSION
 BQ443783.1
 SOURCE
 Mus musculus (house mouse)
 ORGANISM
 Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 REFERENCE
 1 (bases 1 to 732)
 NIH-MGC http://mgi.nci.nih.gov/
 National Institutes of Health, Mammalian Gene Collection (MGC)
 Unpublished
 CONTACT: Robert Strausberg, Ph.D.
 Email: cgabbs@mail.nih.gov
 Tissue Procurement: Dr. James Lin, University of Iowa
 CDNA Library preparation: Dr. M. Bento Soares, University of Iowa
 CDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
 DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
 Clone Distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:
 http://image.llnl.gov
 This clone was contributed by the Brain Molecular Anatomy Project
 (BMAP)

Seq primer: pYX-5,
 Location/Qualifiers
 1. 732
 /organism="Mus musculus"
 /mol_type="mRNA"
 /strain="C57BL/6"
 /db_xref="taxon:10090"
 /clone="IMAGE:5708522"
 /tissue_type="whole brain"
 /lab_host="embryo 15.5 dpc"
 /clone_lib="NIH_BMAP_EVO"
 /notes="Organ: brain; Vector: pYX-Asc; Site 1: EcoR I; Site 2: Not I; The library was constructed according to Bonaldo, Lennon and Soares, Genome Research, 6:791-806, 1996. Denatured mRNA was size fractionated on a 1% agarose gel. First strand cDNA synthesis was primed with an oligo-dT primer containing a Not I site. Double stranded cDNA was size selected according to mRNA size fraction, ligated with EcoR I adaptor, digested with Not I, and then cloned directionally into pYX-Asc vector. The library tag sequence located between the Not I site and the polyA tail, is GTGCGTGGAA. This library was created for the University of Iowa Mouse Brain Molecular Anatomy Project (BMAP): 'Gene Discovery in the Developing Mouse Nervous System', supported by National Institutes of Mental Health (NIMH), Hemin Chin, Ph.D., program coordinator."

BASE COUNT 178 a 240 c 188 g 125 t
 ORIGIN

Query Match 32.9%; Score 601.4; DB 13; Length 732;
 Best Local Similarity 88.8%; Pred. No. 2.2e-122;


```
QY 481 TTGCGACAGCTGCTGGAGAGGTGAAGCGGGCGTGTTCACATGCCGCACTTTATCCCG 540
Db 520 CTGCGGAGTGTGAGAGAGGTCAAGCGTGTGTTCACATGCCGCACTTTATCCCA 579
QY 541 CCGACTGCCAGAGTCTGCTACGGGCGATGATCAGAGTGGAGCGCCGCGCTCACG 600
Db 580 CCAGACTGCCAGAGTCTCTCGCTGGCGATGATGAGGTGGATGAGCTCGGCGCTCACG 639
QY 601 CTAGACACATTCAAGAAACACATATGGTATATAGGGGGGGAAGATGAGCCCGAACAGAG 660
Db 640 CTAGACACATTCAAGAAACACATATGGTATATAGGGGGGGAAGATGAGCCCGAACAGAG 699
QY 661 CAGCCCATTCCTCGCAAGGTGCAGATCCGCTCGCTGCCGCTGCCAGCTGGAGGACATCGACCCC 720
Db 700 CAGCCCATTCACGCAAGGTGCAGATCCGCTCGCTGCCGCTGCCAGCTGGAGGACATCGACCCC 759
QY 721 GAGCTGCTGACAGCATGCACTCACTGGGCTG 752
Db 760 GATGTGTGGACAGCATGCACTCACTGGGCTG 791

RESULT 4
AL538014 1201 bp mRNA linear EST 31-MAY-2003
LOCUS
DEFINITION
AL538014 Homo sapiens FETAL BRAIN Homo sapiens cDNA clone
CS0DF029YC21 5-PRIME, mRNA sequence.
ACCESSION
AL538014.2 GI:31262615
VERSION
EST.
KEYWORDS
SOURCE
ORGANISM
Homo sapiens (human)
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1 (bases 1 to 1201)
Li, W.B., Gruber, C., Jesse, J. and Polayes, D.
Full-length cDNA libraries and normalization
Unpublished
On Feb 13, 2001 this sequence version replaced gi:12801507.
COMMENT
Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 EVRY cedex - France
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
Library was constructed by Life Technologies, a division of
Invitrogen. This sequence belongs to sequence cluster 6094.r For
more information about this cluster, see
http://www.genoscope.cns.fr/
cgi-bin/cluster.cgi?seq=CS0DF029AB11QPI&cluster=6094.r. Contact :
Feng Liang Email : fliang@lifetech.com URL :
http://fulllength.invitrogen.com/ Invitrogen Corporation 1600
Faraday Avenue Genoscope sequence ID : CS0DF029AB11QPI.
Location/Qualifiers
1. 1201
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0DF029YC21"
/tissue_type="FETAL BRAIN"
/dev stage="fetal"
/clone_lib="Homo sapiens FETAL BRAIN"
/notes="Organ: brain; Vector: pCMVSPORT 6; 1st strand cDNA
was primed with a NotI-oligo (dT) primer. Five prime end
enriched, double-strand cDNA was digested with Not I and
cloned into the Not I and EcoRV sites of the pCMVSPORT 6
vector. Library was not normalized."
BASE COUNT 273 a 282 c 364 g 204 t 78 others
ORIGIN
Query Match 32.6%; Score 595.2; DB 9; Length 1201;
Best Local Similarity 95.1%; Pred. No. 5.7e-121;
Matches 645; Conservative 10; Mismatches 18; Indels 5; Gaps 4;
QY 1 ATGAAGTGGAGCGGAGATCGCATCTCGATCTTGAAGCTATTGAGACCCCGCAGCTCTAAAG 60
```

```
Db 266 ATGAAGTGGAGCGGAGATCGCATCTCTGAAGCTATTGAGCACCCCGCAGCTCTAAAG 325
QY 61 CTGACGAGCTTTATGAAAAAATAATTTGTACCTGGTGTAGACAGCTGTCAAGT 120
Db 326 CTGACGAGCTTTATGAAAAAATAATTTGTACCTGGTGTAGACAGCTGTCAAGT 385
QY 121 GGTGAGCTCTTCGACTACCTGCTGAAGAGGGGAGGCTGAGCCCTTAAGGAGGCTCGAAG 180
Db 386 GGTGAGCTCTTCGACTACCTGCTGAAGAGGGGAGGCTGAGCCCTTAAGGAGGCTCGAAG 445
QY 181 TTCTTCGGGAGATCATCTCTGCGCTGGACCTTTGCGCAAGGCACTCCATATGCCACAGG 240
Db 446 TTCTTCGGGAGATCATCTCTGCGCTGGACCTTTGCGCAAGGCACTCCATATGCCACAGG 505
QY 241 GATCTGAAACCTGAAACCTCTCTGCTGACGAGAGAAACAACATCCGATCGACAGCTTT 300
Db 506 GATCTGAAACCTGAAACCTCTCTGCTGACGAGAGAAACAACATCCGATCGACAGCTTT 565
QY 301 GGCATGGGCTCCCTGACAGTGGCGACAGCTTTGGAGACAGCTGTGGGTCCCCCAGC 360
Db 566 GGCATGGGCTCCCTGACAGTGGCGACAGCTTTGGAGACAGCTGTGGGTCCCCCAGC 625
QY 361 TAGGCTCCCGAGGTGATCGGGGGGAGAGTATGACGGCCGGAAGCGGACCTGTGG 420
Db 626 TAGGCTCCCGAGGTGATCGGGGGGAGAGTATGACGGCCGGAAGCGGACCTGTGG 684
QY 421 AGCTCGGCGCTCATCTCTGCTGGTGGGGGCTCTGCGCTTCGACGATGACAAAC 480
Db 685 AGCTCGGCGCTCATCTCTGCTGGTGGGGGCTCTGCGCTTCGACGATGACAAAC 744
QY 481 TTGCGACAGCTGCTCGAAGAGGTGAAGCGGGCGGTTCACATCGCGACCTTTATCCG 540
Db 745 TTGCGACA-CTGCTGGAGAGGTGAAGCGGGCGGTTCACATCGCGACCTTTATCCG 803
QY 541 CCGGACTGCCAGAGTCTGCTACGGGGGATGATCGAGGTGGAGCGCCGCGCTCACG 600
Db 804 CCCSACTCCAGARTCTGCTACGGGGGATGATCGAGGTGGAGCGCCGCGCTCACG 863
QY 601 CTAGAGCATTTCAGAAACACATATGTTATAGGGGCAAGAAATGAGCCCGAACAGAG 660
Db 864 CTAGAGCATTTCAGAAACACATATGTTATAGGGGCAAGAAATGAGCCCGAACAGAG 921
QY 661 CAGCCCATTCCTCGCAAG 678
Db 922 CAG-CCATTCCTCGCAAG 938

RESULT 5
CA315127 723 bp mRNA linear EST 26-NOV-2002
LOCUS
DEFINITION
CA315127 UI-M-FW0-cbi-0-07-0-UI.r1 NIH_BMAP_FW0 Mus musculus cDNA clone
IMAGE: 6810176 5', mRNA sequence.
ACCESSION
CA315127.1 GI:24533251
VERSION
EST.
KEYWORDS
SOURCE
ORGANISM
Mus musculus (house mouse)
Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 723)
NIH-MGC http://mgc.nci.nih.gov/
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished
Contact: Robert Strausberg, Ph.D.
Email: cgaabs@mail.nih.gov
Tissue Procurement: Dr. Jim Lin, University of Iowa
cDNA Library preparation: Dr. M. Bento Soares, University of Iowa
cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
Clone Distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/ILNL at:
http://image.llnl.gov
```


Query Match 32.2%; Score 588.2; DB 13; Length 733;
Best Local Similarity 88.4%; Pred. No. 1.8e-119;
Matches 638; Conservative 0; Mismatches 84; Indels 0; Gaps 0;

QY 285 CCGCATCGCAGACTTTGGCATGGCTCCCTGCGAGTTGGCGACAGCCTGTTGAGACCGAG 344
DB 1 CCGTATTGCGAGACTTTGGCATGGCATCCCTGCGAGTTGGCGACAGCCTGTTGAGACCGAG 60
QY 345 CTGTGGTCCCGCCTACGCTGCGGAGTGTATCCGGGGGAGAGTATGACGCGCG 404
DB 61 CTGCGGATCCACACTATGCTGCGGAGTGTATCCGGGGGAGAGTATGATGCGCG 120
QY 405 GAAGCGGAGTGTGGAGCTGCGGCGTATCCTGTTGCGCTTGTGCTGGGGCTTGCC 464
DB 121 CAAGCGAGATGTGGAGCTGTGTGTGATCCTGTTGCGCTTGTGCTGGGGCTTGCC 180
QY 465 CTTCCGAGTACAACTTGGCAGAGCTGCTGGAGAGGTGAAGCGGGCGTGTGTCACAT 524
DB 181 TTTTGTATGATGACAACTGCGGCAAGTTGCTGGAGAGGTCAAGCGTGTGTGTTCCACAT 240
QY 525 GCGCAGCTTTATCCCGCCGAGTCCGAGTCTGCTACGGGGCATGATCGAGTGGAGCG 584
DB 241 GCCACACTTTATCCACAGACTGCCAGAGTCTCCTGCTGGCATGATTCAGGTGGATGC 300
QY 585 CGCAGCGCTCAGCTAGAGCACATTCAGAACACATATGTTATAGGGGCAAGAA 644
DB 301 AGCTCGGGCGCTCAGCTAGAGCACATTCAGAACACATATGTTATAGGTGGCAAGAA 360
QY 645 TGAGCCGCAACGAGCAGCGCCATTCCTCGCAAGGTGCAGATCCGCTCGCTGCCAGCCT 704
DB 361 TGAGCCGAGCGCGAAGCAGCCATCCACGCAAGGTGCAGATCCGCTCACTACCCAGCTT 420
QY 705 GGAGGACATGACCCCGAGCTGTGAGAGCATGCTACTCTGGCTGTCTTCGAGACCG 764
DB 421 GGAAGACATTCAGCTGTATGTTGAGAGCATGCTACTCTGGCTGTCTTCGAGACCG 480
QY 765 CAAAGCTGCTGAGGACCTGCTGCTCGAGAGGAGAACCCAGGAGAGATGATTTACTT 824
DB 481 CAAAGCTGCTGAGGATCTGCTATCTGAGAGGAGAACTCAGGAAAGATGATTTATTT 540
QY 825 CCTCTCTGACCGGAAAGAGTATCCCGAGCGAGGAGATGAGGACCTGCCCCCGCG 884
DB 541 CCTCTCTGATCGGAAAGAACGTTATCCAAAGCATGAGGATGAGGACCTGCCCCCGAG 600
QY 885 GAACGAGTAGACCTTCCCGGAGCGTGTGAGATCTCCCGATGTGAAACCGGACGCA 944
DB 601 GAATGAGTAGACCTTCCCGGAGCGTGTGAGATCTCCCGATGTGAAACCGGCAATGCAA 660
QY 945 GCGGCGCCAGACGCAATCCATGGAGTGTCTCAGCGTGAACGCGCGCTCCCGGT 1004
DB 661 GCGGCGACCTGAGCGCAAGTCCATGGAAAGTGTCTCAGTGTGACAGATGTGCTCCAGNT 720
QY 1005 GC 1006
DB 721 GC 722

RESULT 7
BG395625
LOCUS
DEFINITION 733 bp mRNA linear EST 12-MAR-2001
602458251F1 NIH_MGC_16 Homo sapiens cDNA clone IMAGE:4580724 5',
mRNA sequence.
ACCESSION BG395625
VERSION BG395625.1 GI:13289073
KEYWORDS EST.
SOURCE Homo sapiens
ORGANISM Homo sapiens (human)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1 (bases 1 to 733)
AUTHORS NIH-MGC <http://mgc.nci.nih.gov/>.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished

COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: ATCC
cDNA Library Preparation: Ling Hong/Rubin Laboratory
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone Distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at: image.llnl.gov
Plate: LLCM1300 row: e column: 13
High quality sequence stop: 719.

FEATURES
Location/Qualifiers
1..733
Source
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:4580724"
/tissue_type="retinoblastoma"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH MGC 16"
/note="Organ: eye; Vector: pOTB7; Site_1: XhoI; Site_2:
EcoRI; cDNA made by oligo-dT priming. Directionally
cloned into EcoRI/XhoI sites using the following 5'
adaptor: GGCACGAG(G). Library constructed by Ling Hong
in the laboratory of Gerald M. Rubin (University of
California, Berkeley) using ZAP-cDNA synthesis kit
(Stratagene) and Superscript II RT (Life Technologies).
Note: this is a NIH MGC library."

BASE COUNT 162 a 238 c 214 g 119 t
ORIGIN

Query Match 32.0%; Score 584.8; DB 10; Length 733;
Best Local Similarity 99.7%; Pred. No. 1e-118; 2; Indels 0; Gaps 0;
Matches 586; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1179 GGAGAGCCCGGTGGCAGCGCCCAACCCCAACCCCGCTCCAGCCCCCAGCGGTGGAGGGGT 1238
DB 2 GGAGAGCCCGGTGGCAGCGCCCAACCCCAACCCCGCTCCAGCCCCCAGCGGTGGAGGGGT 61
QY 1239 GCCTTGGAGGGCGGGCTCACTCCATCAAGAACAGCTTTCTGGGCTACCCGGCTTCCA 1298
DB 62 GCCTTGGAGGGCGGGCTCACTCCATCAAGAACAGCTTTCTGGGCTACCCGGCTTCCA 121
QY 1299 CCGCCGGAACTGCAAGTTCCGACGCGGAGAGATGTCACACCTGACACGAGATCGTC 1358
DB 122 CCGCCGGAACTGCAAGTTCCGACGCGGAGAGATGTCACACCTGACACGAGATCGTC 181
QY 1359 CCAGAGCTGGCGAAGTCTCTGGTTTGGGAATTCATCAGCCTGGAGAGAGAGCA 1418
DB 182 CCAGAGCTGGCGAAGTCTCTGGTTTGGGAATTCATCAGCCTGGAGAGAGAGCA 241
QY 1419 GATCTTGTGGTTCATCAAGACAACTCTGAGCTCCATCAGGCTGACATCGTGACGC 1478
DB 242 GATCTTGTGGTTCATCAAGACAACTCTGAGCTCCATCAGGCTGACATCGTGACGC 301
QY 1479 CTTCTGTTCGATTCCTCAGCCAGCTCTCAGCCAGCGTCTCTCCAAACGAGCTTCCGGGCCGA 1538
DB 302 CTTCTGTTCGATTCCTCAGCTCTCAGCCAGCGTCTCTCCAAACGAGCTTCCGGGCCGA 361
QY 1539 GTACAAAGCCACGGGGGGCCAGCGCTGTTCAGAAAGCCGGTCAAGTTCAGGTTGATAT 1598
DB 362 GTACAAAGCCACGGGGGGCCAGCGCTGTTCAGAAAGCCGGTCAAGTTCAGGTTGATAT 421
QY 1599 CACTTACAGGAGGTGGGAGGCGCAGAGGAGGAGGATCTACTCCGTTCACCTTCAC 1658
DB 422 CACTTACAGGAGGTGGGAGGCGCAGAGGAGGAGGATCTACTCCGTTCACCTTCAC 481
QY 1659 CTTGCTCTCAGGCGCCAGCGCTCGCTTCAAGAGGGTGGAGACCATCCAGGCCAGCT 1718
DB 482 CTTGCTCTCAGGCGCCAGCGCTCGCTTCAAGAGGGTGGAGACCATCCAGGCCAGCT 541
QY 1719 GCTGAGCACACAGCCCGCTGCGGCCAGCAGCTTGTGACAGACCAAC 1766
DB 542 GCTGAGCACACAGCCCGCTGCGGCCAGCAGCTTGTGACAGAAACCC 589

1996. Denatured mRNA was size fractionated on a 1% agarose gel. First strand cDNA synthesis was primed with an oligo-dT primer containing a Not I site. Double stranded cDNA was size selected according to mRNA size fraction, ligated with EcoR I adaptor, digested with Not I, and then cloned directionally into pYX-Asc vector. The library tag sequence located between the Not I site and the polyA tail is GTGGTGGAA. This library was created for the University of Iowa Mouse Brain Molecular Anatomy Project (BMAP): 'Gene Discovery in the Developing Mouse Nervous System', supported by National Institutes of Mental Health (NIMH), Hemin Chin, Ph.D., program coordinator."

BASE COUNT 167 a 211 c 193 g 139 t 2 others
ORIGIN
Query Match 31.4%; Score 574.2; DB 13; Length 712;
Best Local Similarity 88.6%; Pred. No. 2.2e-116;
Matches 632; Conservative 0; Mismatches 80; Indels 1; Gaps 1;
QY 463 CCCTTCGACGATGACAACTTTCGACAGCTGCTCGAGAGGTGAAGCGGGCGTGTTCAC 522
DB 1 CTTTGTGATGATGACAACTTTCGCGAGTGTCTCGAGAGGTCAAGCGTGTGTTCAC 60
QY 523 ATGCGGCACATTTATCCCGCCGACTGCGAGAGTCTGCTACGGGGCATGATCGAGGTGAC 582
DB 61 ATGCCACATTTATCCCGCCGACTGCGAGAGTCTGCTGCGTGGCATGATTGAGTGGAT 120
QY 583 GCGCGACGCGCTCAGCTAGAGCATTTCAGAAACATATGTTATATAGGGGCAAG 642
DB 121 GCAGCTCGGCGCTCAGCTAGAGCATTTCAGAAACATATGTTATATAGTGGCAAG 180
QY 643 AATGAGCCGAAACAGAGCAGCCATTCCTCGAAGGTGCAGATCCGCTCGTCCCGACG 702
DB 181 AATGAGCCGAGCCGAGAGCCATTCAGAGGTGCAGATCGCTCACTACCGAGC 240
QY 703 CTGAGAGCATTCGACCCGAGCTGTGGAAGCATGACATCTCATGGCTGCTTCGAGAC 762
DB 241 TTGGAAGACATTTGACCTGTATGTTGAGAGCATGACATCACTGGGCTGCTTCGAGAC 300
QY 763 CGCAACAGCTGCTGAGGACCTGCTGTCGAGAGGAGAGAACAGAGAGATGATTAC 822
DB 301 CGCAACAGCTGCTGAGGATCTGCTATCTGAGAGGAGAGATCAGGAAAGATGATTAT 360
QY 823 TTCTCTCTCTGAGCCGGAAGAGGTATCCGAGCCAGGAGATGAGACCTGCCCGCC 882
DB 361 TTCTCTCTCTGATCGGAAGAGCGTATCCAGCCATGAGATGAGACCTGCCCGCC 420
QY 883 CGGAACGATAGACCTTCCCGGAAGCGGTGTGACATCCCGCATGCTGAACCGGCGCG 942
DB 421 AGGAATGAGATAGACCTTCCCGGAAGCGGTGTGATTTCCCGCATGCTGAACCGGCGATGGC 480
QY 943 AAGCGGCGCCAGAACCAATCCATGAGGTGCTCAGCGTGAGGAGCGGCTCCCGC 1002
DB 481 AAGCGGCGACTGAGCGCAAGTCCATGAAGTGTCTAGTGTGACAGATGTGGCTCCCA 540
QY 1003 GTGCTCGCGCGCGGCAATGAGATGCCAGACGCGGAGAGGTCTCGGTCCATCAGC 1062
DB 541 GTGCTCGACGAGAGCATTTGATGGCCAGCATGCGCAGAGATCTCGATCCATCAGT 600
QY 1063 GGTGCTCTCAGGCTTTTCCACAGGCCACTACAGAGCCCGCGGTGACCCCTCACCCC 1122
DB 601 GGTGCTCTCAGGCTTTTTCACAGTCCACTACGAGTCTCTCGGCTGACCCCTCACCCC 660
QY 1123 TCACCAAGGGGCACTCCCTCCCGACCCCAAGGGGACCTGTCCACAGCC 1175
DB 661 TCACCAAGGGGTAGT-CCCTTCTTACCCCAAGGGGAGCGCTGTCCACAGCC 712

RESULT 10
BF222739/c 740 bp mRNA linear EST 30-MAR-2001
LOCUS 7q32f02.x1 NCI CGAP GC6 Homo sapiens cDNA clone IMAGE:369987 3'
DEFINITION similar to TR:O60843 O60843 PUTATIVE SERINE/THREONINE PROTEIN

KINASE i, mRNA sequence.
BF222739
VERSION BF222739.1 GI:11129830
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Mammalia; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Eukaryota; Eukarya; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 740)
AUTHORS NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.
TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP), Tumor Gene Index
JOURNAL Unpublished
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs@mail.nih.gov
Tissue Procurement: Christopher A. Moskaluk, M.D., Ph.D., Michael R. Emmert-Buck, M.D., Ph.D.
cDNA Library Preparation: M. Bento Soares, Ph.D., M. Fatima Bonaldo, Ph.D.
cDNA Library Arrayed by: Greg Lennon, Ph.D.
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be found through the I.M.A.G.E. Consortium/ILNL, send email to: info@image.llnl.gov
High quality sequence stop: 492.

FEATURES
Location/Qualifiers
1..740
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:369987"
/tissue_type="pooled germ cell tumors"
/lab_host="DH10B"
/clone_lib="NCI CGAP GC6"
/note="Vector: pRT3D-Pac (Pharmacia) with a modified polylinker; Site 1: Not I; Site 2: Eco RI; Plasmid DNA from the normalized library NCI CGAP GC4 was prepared, and ss circles were made in vitro. Following HAP purification, this DNA was used as tracer in a subtractive hybridization reaction. The driver was PCR-amplified cDNAs from a pool of 5,000 clones made from the same library (cloneIDs 1257096-1259631, 1469064-1470983, and 1475592-1476743). Subtraction by Bento Soares and M. Fatima Bonaldo."

BASE COUNT 125 a 206 c 241 g 163 t 5 others
ORIGIN

Query Match 30.8%; Score 562.2; DB 10; Length 740;
Best Local Similarity 97.8%; Pred. No. 9.9e-114;
Matches 578; Conservative 0; Mismatches 12; Indels 1; Gaps 1;

QY 1221 CCCACGCTCGGA-GGGGTGCGCTGGAGGGCGGGCTCAACTCCATCAAGAACAGCTTTC 1279
DB 721 CCCACGCTCGNAGGGGTGCCNTGGAGGGCGGGCTCAACTCCATCAAGAACAGCTTTC 662
QY 1280 TGGGCTCACCCCGTTCACCCCGGAAACTCGAAGTTCGACGCGCGAGGAGATGCCA 1339
DB 661 TGGGCTCACCCCGTTCACCCCGGAACTGCAAGTTTCGACGCGCGAGGAGATGCCA 602
QY 1340 ACTGACACAGAGTCTGCCAGAGCTGGCAAGAGTCTGTTGGGAACTTCATCA 1399
DB 601 ACTGACACAGAGTCTGCCAGAGCTGGCAAGAGTCTGTTGGGAACTTCATCA 542
QY 1400 GCTGAGAGAGGAGGAGAGATCTTCGTGGTTCATCAAGACAAACCTCTGAGCTCCATCA 1459
DB 541 GCTGAGAGAGGAGGAGAGATCTTCGTGGTTCATCAAGACAAACCTCTGAGCTCCATCA 482
QY 1460 AGGCTGACATCGTGACAGCTTCTCTGTCCAGTCTCCAGTCTCCAGCCACAGGCTCATCCC 1519
DB 481 AGGCTGACATCGTGACAGCTTCTCTGTCCAGTCTCCAGTCTCCAGCCACAGGCTCATCCC 422
QY 1520 AAACGAGCTTCGGGGCGAGTACAGGCCACGGGGGGCGAGCGGTGTTCCAGAGCCCGG 1579
DB 421 AAACGAGCTTCGGGGCGAGTACAGGCCACGGGGGGCGAGCGGTGTTCCAGAGCCCGG 362

```

QY 1580 TCAGTTCCAGTTGATATCATCTACCGAGGGTGGGAGGCGCAGAGAGAACGGCA 1639
Db |||||
QY 361 TCAAGTTCAGGTTGATATCATCTACCGAGGGTGGGAGGCGCAGAGAGAACGGCA 302
Db |||||
QY 1640 TCTACTCGTTCACCTTTCACCTTCTCAGGCCCGCCGCTTCAAGAGGGTGGTGG 1699
Db |||||
QY 301 TCTACTCGTTCACCTTTCACCTTCTCAGGCCCGCCGCTTCAAGAGGGTGGTGG 242
Db |||||
QY 1700 AGACATCCAGGCCAGCTGTGTGAGCACACAGACCCCGCTTCCAAATGTGGA 1759
Db |||||
QY 241 AGACATCCAGGCCAGCTGTGTGAGCACACAGACCCCGCTTCCAAATGTGGA 182
Db |||||
QY 1760 ACACCACTAATCTATGGAATATGACGGGGCGCTTCCAAATGTGGA 1810
Db |||||
QY 181 ACACCACTAATCTATGGAATATGACGGGGCGCTTCCAAATGTGGA 131
Db |||||

RESULT 11
LOCUS BE251924 563 bp mRNA linear EST 13-JUL-2000
DEFINITION 601107590F1 NIH_MGC_16 Homo sapiens cDNA clone IMAGE:3343822 5',
mRNA sequence.
ACCESSION BE251924
VERSION BE251924.1 GI:9122058
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
TITLE NIH-MGC http://mgi.nci.nih.gov/
JOURNAL National Institutes of Health, Mammalian Gene Collection (MGC)
COMMENT Unpublished
Contact: Robert Strausberg, Ph.D.
Email: cgabbs-r@mail.nih.gov
Tissue Procurement: ATCC
CDNA Library Preparation: Ling Hong/Rubin Laboratory
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LML)
DNA Sequencing by: Incyte Genomics, Inc.
Clone Distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at: image.llnl.gov
Plate: L1CM129 row: c column: 23
High quality sequence stop: 563.
FEATURES
Location/Qualifiers
source
1..563
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:3343822"
/tissue_type="retinoblastoma"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH_MGC_16"
/note="Organ: eye; Vector: pOTB7; Site 1: XhoI; Site 2:
EcoRI; cDNA made by oligo-dT priming. Directionally
cloned into EcoRI/XhoI sites using the following 5'
adaptor: GGCACGAG(G). Library constructed by Ling Hong
in the laboratory of Gerald M. Rubin (University of
California, Berkeley) using ZAP-cDNA synthesis kit
(Stratagene) and Superscript II RT (Life Technologies).
Note: this is a NIH_MGC Library."
BASE COUNT 122 a 190 c 158 g 93 t
ORIGIN

Query Match 30.6%; Score 559.8; DB 10; Length 563;
Best Local Similarity 99.6%; Pred. No. 3.1e-113;
Matches 561; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1198 CCCAACCCCGCCCGCTCCAGCCCGCCAGCGTCGAGGGGTGCCCTGAGGGCGCGCTC 1257
Db |||||
QY 1258 AACTCCATCAAGACAGCTTTCTGGGCTACCCCGCTTCCACCGCCGGAACCTGCAAGTT 1317
Db |||||

```

```

Db 61 AACTCCATCAAGAACAGCTTTCTGGGCTCACCCCGCTTCCACCGCCGGAACCTGCAAGTT 120
QY 1318 CCAGACGCCGAGGAGAGATGTCCAACTGACACAGAGTCGTCGCCAGAGCTGGCGAAGAAG 1377
Db |||||
QY 121 CCAGACGCCGAGGAGAGATGTCCAACTGACACAGAGTCGTCGCCAGAGCTGGCGAAGAAG 180
Db |||||
QY 1378 TCCTGTGTTGGGAACTTCATCAGCCTTGGAGAGGAGAGATCTTCGTGTGTCATCAAA 1437
Db |||||
QY 181 TCCTGTGTTGGGAACTTCATCAGCCTTGGAGAGGAGAGATCTTCGTGTGTCATCAAA 240
Db |||||
QY 1438 GACAAACCTCTGAGCTCCATCAAGGCTGACATCGTCAGCGCTTCCCTGTGTCATTCACAGT 1497
Db |||||
QY 241 GACAAACCTCTGAGCTCCATCAAGGCTGACATCGTCAGCGCTTCCCTGTGTCATTCACAGT 300
Db |||||
QY 1498 CTGAGCCACAGCGCTCATCTCCAAACGAGCTTCCGGGCGGAGTACAAGGCCACGCGGGGG 1557
Db |||||
QY 301 CTGAGCCACAGCGCTCATCTCCAAACGAGCTTCCGGGCGGAGTACAAGGCCACGCGGGGG 360
Db |||||
QY 1558 CGAGCGGTGTTCCAGAACGCGGTCAAGTTCAGGTTGATATCACTACACGAGGGTGGG 1617
Db |||||
QY 361 CCAGCGGTGTTCCAGAACGCGGTCAAGTTCAGGTTGATATCACTACACGAGGGTGGG 420
Db |||||
QY 1618 GAGCGCAGAGGAGAGACGSCATCTACTCGTCACCTTCACTCAGCGCTTCTCAGCGCCCGCAGC 1677
Db |||||
QY 421 GAGCGCAGAGGAGAGACGSCATCTACTCGTCACCTTCACTCAGCGCTTCTCAGCGCCCGCAGC 480
Db |||||
QY 1678 CGTCGCTTCAAGAGGGTGGTGGAGAGCATTCAGCGCCCGAGTGTGAGCACACACGACCCG 1737
Db |||||
QY 481 CGTCGCTTCAAGAGGGTGGTGGAGAGCATTCAGCGCCCGAGTGTGAGCACACACGACCCG 540
Db |||||
QY 1738 CTTGCGGCCCGCAGCACTTGTGAGA 1760
Db |||||
QY 541 CTTGCGTGCAGCACTTGTGAGA 563
Db |||||

RESULT 12
LOCUS BQ43644 693 bp mRNA linear EST 29-MAY-2002
DEFINITION UI-M-EWO-bxh-n-01-0-UI.r1 NIH_BMAP_EWO Mus musculus cDNA clone
IMAGE:5708352 5', mRNA sequence.
ACCESSION BQ43644
VERSION BQ43644.1 GI:21246756
KEYWORDS EST.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
TITLE NIH-MGC http://mgi.nci.nih.gov/
JOURNAL National Institutes of Health, Mammalian Gene Collection (MGC)
COMMENT Unpublished
Contact: Robert Strausberg, Ph.D.
Email: cgabbs-r@mail.nih.gov
Tissue Procurement: Dr. James Lin, University of Iowa
CDNA Library Preparation: Dr. M. Bento Soares, University of Iowa
CDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
Clone Distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
This clone was contributed by the Brain Molecular Anatomy Project
(BMAP)
Seq primer: pYX-5.
Location/Qualifiers
1..693
/organism="Mus musculus"
/mol_type="mRNA"
/db_xref="taxon:10090"
/clone="IMAGE:5708352"
/tissue_type="whole brain"
/dev_stage="embryo 15.5 dpc"

```

```
/lab_host="DH10B (T1 phage resistant)"
/clone_lib="NIH_BMAP_EW0"
/site="Organ: brain; Vector: pYX-Asc; Site 1: EcoR I;
Site 2: Not I; The library was constructed according to
Bonaldo, Lennon and Soares, Genome Research, 6:791-806,
1996. Denatured mRNA was size fractionated on a 1% agarose
gel. First strand cDNA synthesis was primed with an
oligo-dT primer containing a Not I site. Double stranded
cDNA was size selected according to mRNA size fraction,
ligated with EcoR I adaptor, digested with Not I, and then
cloned directionally into pYX-Asc vector. The library tag
sequence located between the Not I site and the polyA tail
, is GTGGCTGGA. This library was created for the
University of Iowa Mouse Brain Molecular Anatomy Project
(BMAP): 'Gene Discovery in the Developing Mouse Nervous
System', supported by National Institutes of Mental Health
(NIMH), Hemin Chin, Ph.D., program coordinator."

BASE COUNT      168 a      179 c      198 g      147 t
ORIGIN
Query Match      30.0%; Score 548.2; DB 13; Length 693;
Best Local Similarity 87.8%; Pred. No. 1.2e-110;
Matches 609; Conservative 0; Mismatches 84; Indels 1; Gaps 1;

QY 207 GGACTTCTGCCACGACCTCCATATGCCACAGGATCTGAACACCTGAAACCTCTGCT 266
Db 1 GGACTTCTGCCACGACCTCCATATGCCATAGAGACTTGAAGCCAGAGAACCTGCTGCT 60

QY 267 GGACGAGAGAACACATCCGATCCGACACTTGGCATGGCTCCCTGCAGGTGGCGA 326
Db 61 AGATGAGAGAAC-ACATCCGATTTGCAGACTTTGGCATGGCATCCCTGCAGGTGGCGA 119

QY 327 CAGCCTCTGGAGACCACTGTGGTCCCCCCTACTACGCTGCCCGAGGTGATCCGGGG 386
Db 120 CAGCCTCTGGAGACCACTGTGGATCTCCACATATGCTCTGCGGAGTGTTCGGGG 179

QY 387 GGAAGATATGACGGCCGGAAGCGGAGTGTGAGCTGCGGGCTCATCTGTTCGCTT 446
Db 180 CGAAGATATGATGGCCGCAAGGAGAGTGTGAGCTGTGCTGTATCTGTTCGCTT 239

QY 447 GCTGTTGGGGCTCTGCTTCCGATGACACTTGGACAGCTCTGAGAGAGTGA 506
Db 240 GCTGTTGGGGCTCTGCTTTCATGATGACACCTTCGCGGAGTGTGAGAGAGTGA 299

QY 507 GCGGGGCTGTTCACATGCGGACCTTTATCCGCGCCGACTGCCAGAGTCTGTACGGGG 566
Db 300 GCGTGGTGTGTTCACATGCGGACCTTTATCCGACAGACTGCCAGAGTCTGTGCTGG 359

QY 567 CATGATCGAGGTGAGCGCGGACGCGCTTACGCTAGAGACATTTGAGAGAACATATG 626
Db 360 CATGATCGAGGTGAGTGCAGCTCGGCGCTTACGCTAGAGACATTTGAGAGAACATATG 419

QY 627 GTATATAGGGGGAAGATGAGCCGACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAT 686
Db 420 GTATATAGGGTGGCAAGATGAGCCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAT 479

QY 687 CCGCTCGCTCCGAGCTTGAGAGACATCGACCCCGAGCTGTGAGAGAGATGCACTCACT 746
Db 480 CCGCTCACTACCGAGCTTGGAGACATGAGCCCTGATGTGTGGAGAGAGATGCACTCACT 539

QY 747 GGGTGTCTTCGAGAGACCGGACAAAGAGTGTGTCGAGAGACCTGCTGTGAGAGAGAGACCA 806
Db 540 GGGTGTCTTCGAGAGACCGGACAAAGAGTGTGTCGAGAGATCTGCTATCTGAGGAGAGATCA 599

QY 807 GGAGAGATGATTTACTTCTCTCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 866
Db 600 GGAGAGATGATTTACTTCTCTCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 659

QY 867 TGAGGACCTGCCCGCCCGGAGACGAGATAGACCTT 900
Db 660 TGAGGACCTGCCCGCCCGGAGATGAGATAGACCTT 693
```

```
RESULT 13
BE259121
LOCUS      BE259121
DEFINITION 601107773F1 NIH_MGC_16 Homo sapiens cDNA clone IMAGE:3343795 5',
            mRNA sequence.
ACCESSION  BE259121
VERSION    BE259121.1 GI:9129619
KEYWORDS   EST.
SOURCE     Homo sapiens (human)
ORGANISM   Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE  1 (bases 1 to 589)
            NIH-MGC http://mgc.nci.nih.gov/.
AUTHORS   National Institutes of Health, Mammalian Gene Collection (MGC)
TITLE     Unpublished
JOURNAL    Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-remail.nih.gov
            Tissue Procurement: ATCC
            cDNA Library Preparation: Ling Hong/Rubin Laboratory
            cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
            DNA Sequencing by: Incyte Genomics, Inc.
            Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at: image.llnl.gov
            Plate: LLCM129 row: b column: 20.

FEATURES             Location/Qualifiers
     1..589
         /organism="Homo sapiens"
         /mol_type="mRNA"
         /db_xref="taxon:9606"
         /clone="IMAGE:3343795"
         /tissue_type="retinoblastoma"
         /lab_host="DH10B (phage-resistant)"
         /clone_lib="NIH MGC 16"
         /note="Organ: eye; Vector: pOTB7; Site 1: XhoI; Site 2:
         EcoRI; cDNA made by oligo-dT priming. Directionally
         cloned into EcoRI/XhoI sites using the following 5'
         adaptor: GGCACGAG(G). Library constructed by Ling Hong
         in the laboratory of Gerald M. Rubin (University of
         California, Berkeley) using ZAP-cDNA synthesis kit
         (Stratagene) and Superscript II RT (Life Technologies).
         Note: this is a NIH_MGC Library."
BASE COUNT      125 a      190 c      177 g      97 t
ORIGIN
Query Match      28.2%; Score 515; DB 10; Length 589;
Best Local Similarity 99.0%; Pred. No. 2.5e-103;
Matches 518; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1174 CCAAGGAGAGCCCGCTGGCACGCGCCCAACCCACGCCCCCTCCAGCCCCAGCGTCGGA 1233
Db 31 CGAGGAGAGAGCCCGCTGGCACGCGCCCAACCCACGCCCCCTCCAGCCCCAGCGTCGGA 90

QY 1234 GGGGTGCCCTGGAGGCGCGGCTCAACTCCATCAGAACAGCTTTCTGGGCTCACCCCGC 1293
Db 91 GGGGTGCCCTGGAGGCGCGGCTCAGCGCCATCAAGAACAGCTTTCTGGGCTCACCCCGC 150

QY 1294 TTCCACCGCGGAAACTGCAAGTTCGACGCGCGGAGGAGATGTCCAACTGACACAGAG 1353
Db 151 TTCCACCGCGGAAACTGCAAGTTCGACGCGCGGAGGAGATGTCCAACTGACACAGAG 210

QY 1354 TCCTCCCCAGAGCTGGCGAAGATCTCTGGTTTGGGAATTCATCAGCCTGAGAGAGAG 1413
Db 211 TCCTCCCCAGAGCTGGCGAAGATCTCTGGTTTGGGAATTCATCAGCCTGAGAGAGAG 270

QY 1414 GAGCAGATCTTCTGGTGCATCAAGACAACTCTGAGCTCCATCAAGCTGCATCTCGTG 1473
Db 271 GAGCAGATCTTCTGGTGCATCAAGACAACTCTGAGCTCCATCAAGCTGCATCTCGTG 330

QY 1474 CAGGCTTCTCTGTCGATTCCTCAGTCTCAGCCAAGCTCATCTCCCAACAGAGCTTCCGG 1533
Db 331 CAGGCTTCTCTGTCGATTCCTCAGTCTCAGCCAAGCTCATCTCCCAACAGAGCTTCCGG 390
```

QY 1534 GCGAGTACACGCGCCAGCGGGGGGCGCCAGCGGTGTTCCAGAGCCGGTCAAGTTCAGGTT 1593
 Db 391 GCGAGTACACGCGCCAGCGGGGGGCGCCAGCGGTGTTCCAGAGCCGGTCAAGTTCAGGTT 450
 QY 1594 GATATCACTACACGAGGGTGGGAGCGCGAGAGGAGGAGCGGCATCTACTCCGTAC 1653
 Db 451 GATATCACTACACGAGGGTGGGAGCGCGAGAGGAGGAGCGGCATCTACTCCGTAC 510
 QY 1654 TTCACCTGCTCTCAGGCGCCAGCGCGTCCGTTCAAGAGGGTGG 1696
 Db 511 TTCACCTGCTCTCAGGCGCCAGCGCGTCCGTTCAAGAGGGTGG 553

RESULT 14
 CA319314 699 bp mRNA linear EST 26-NOV-2002
 LOCUS UI-M-FW0-Cbz-i-20-0-UI.r1 NIH_BMAP_FW0 Mus musculus cDNA clone
 DEFINITION IMAGE: 6816573 5', mRNA sequence.

ACCESSION CA319314
 VERSION CA319314.1 GI:24537438
 KEYWORDS EST.
 SOURCE Mus musculus (house mouse)
 ORGANISM Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sclurognathi; Muridae; Murinae; Mus.
 1 (bases 1 to 699)
 NIH-MGC <http://mgc.nci.nih.gov/>.
 National Institutes of Health, Mammalian Gene Collection (MGC)
 Unpublished
 Contact: Robert Strausberg, Ph.D.
 Email: cgapbs-r@mail.nih.gov
 Tissue Procurement: Dr. Jim Lin, University of Iowa
 cDNA Library preparation: Dr. M. Bento Soares, University of Iowa
 DNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
 DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
 Clone Distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
 This clone was contributed by the Brain Molecular Anatomy Project
 (BMAP)

Seq primer: pYX-5,
 Location/Qualifiers
 1..699
 /organism="Mus musculus"
 /mol_type="mRNA"
 /strain="C57BL/6"
 /db_xref="taxon:10090"
 /clone="IMAGE: 6816573"
 /tissue_type="whole brain"
 /dev_stages="embryo 13.5,14.5,16.5,17.5dpc"
 /lab_host="DH10B (T1 phage resistant)"
 /clone_lib="NIH_BMAP_FW0"
 /note="Organ: Brain; Vector: pYX-Asc; Site 1: EcoR I;
 Site 2: Not I; The library was constructed according
 Bonaldo, Lennon and Soares, Genome Research, 6:791-806,
 1996. Denatured RNA was size fractionated on a 1% agarose
 gel. First strand cDNA synthesis was primed with oligo-dT
 primer containing a Not I site. Double strand cDNA was
 size selected according to mRNA size fraction, ligated
 with EcoR I adaptor, digested with NotI and then cloned
 directionally into pYX-Asc vector. The library tag
 sequence located between the Not I site and the polyA tail
 is AGCGAGACAG. This library was created for the University
 Iowa Brain Anatomy Project (BMAP). 'Gene Discovery in the
 Developing Mouse Nervous System', supported by National
 Institute of Mental Health (NIMH), Hemin Chin, Ph.D.,
 Program coordinator."

BASE COUNT 169 a 237 c 182 g 108 t 3 others
 ORIGIN

Query Match 27.2%; Score 496.8; DB 14; Length 699;
 Best Local Similarity 86.9%; Pred. No. 2.7e-99;
 Matches 546; Conservative 0; Mismatches 82; Indels 0; Gaps 0;

1150 CCAGAGGGGACACCTGTCCACACGCCAAAGAGAGCCGGTGGCAAGCCCAACCCACG 1209
 Db 1 CCAGAGGGGACACCTGTCCACACGCCAAAGAGAGCCGGTGGCAAGCCCAACCCACA 60
 QY 1210 CCCCCTGTCAGACCCACCGCTGGAGGGTGCCTTGGAGGGCGGGCTCAACTCCATCAAG 1269
 Db 61 CCACATCCAGCCCTAGTGTGGAGGAGTGCCCTGGCGGACACGACTGAACTCCATCAAG 120
 QY 1270 AACAGCTTTCTGGGCTCACCCCGTTCACCGCGGGAACCTGCAAGTTCGACGCCGGAG 1329
 Db 121 AACAGCTTTCTGGGCTCACCTCGAATCCACGCGGGAACCTCCAAAGTTCACACCCAGAG 180
 QY 1330 GAGATGTCCAACCTGACACACAGAGTCCCTCCACAGAGCTGGCGAAGAAAGTCTCTGTTGGG 1389
 Db 181 GAGATGTCCAACCTGACACCCAGAGTCTCTCCAGAGCTGGCCAAAGAAATCGTGGTTCGGG 240
 QY 1390 AACTTCATCAGCTTGGAGAGAGGAGGAGCAGATCTTCGTGTGTCATCAAGACAAACCTCTG 1449
 Db 241 AACTTCATCAACCTGGAGAGAGGAGGAGCAGATCTTTGTGCTGATCAAGGACAAGCCCTG 300
 QY 1450 AGCTCCATCAAGGCTGACATCGTCAGCGCTTCTGTGATTCCTCAGTCTTCAGCCACAGC 1509
 Db 301 AGCTCCATCAAGGCTGACATCGTTCATGCTTCTGTGATTCCTCAGTCTTCAGCCACAGC 360
 QY 1510 GTCAATCTCCCAAAACGAGCTTCCGGGGCGAGTACAAGGCCACGGGGGGCCAGCGGTGTT 1569
 Db 361 GTATATTTCCACAGACAAGCTTCAGGGCTGAATACAGGCCACAGGGGGCCACAGCAGTGTTC 420
 QY 1570 CAGAAGCGGTCAAGTTCAGGTTGATATCACTACACGAGGGTGGGGAGGCGCAGAAG 1629
 Db 421 CAGAAGCGGTCAAGTTCAGGTTGATATCACTACACGAGGGTGGGGAGGCGCAGAAG 480
 QY 1630 GAGAACGGCATCTACTCCGTCACCTTCACCTGCTCTCAGGCCCCCAGCGCTCGTTCAG 1689
 Db 481 GAGAAATGGCATCTACTCAGTCACATTCACCTTACTCTCAGGCCCCCAGTCCGCGTTCAG 540
 QY 1690 AGGTGTGTGAGAGCACTCCAGGCGCCAGCTGTGAGCAGACACACCGCCCTCGCGGCCAG 1749
 Db 541 AGGTGTGTGAGAGCACTCCAGGCGCCAGCTGTGAGCAGCACTTAAGCACCATTACAGCCATCAGCCAG 600
 QY 1750 CACTTGTGACAGCACCACCTAACTGTATGG 1777
 Db 601 CACTTGTGACAGCAGAGAAACGGGCGAG 628

RESULT 15
 BI825755
 LOCUS 603072158F1 NIH_MGC_119 Homo sapiens cDNA clone IMAGE:5163999 5',
 DEFINITION mRNA sequence.
 ACCESSION BI825755
 VERSION BI825755.1 GI:15937305
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 1 (bases 1 to 727)
 NIH-MGC <http://mgc.nci.nih.gov/>.
 National Institutes of Health, Mammalian Gene Collection (MGC)
 Unpublished
 Contact: Robert Strausberg, Ph.D.
 Email: cgapbs-r@mail.nih.gov
 Tissue Procurement: Life Technologies, Inc.
 cDNA Library Preparation: Life Technologies, Inc.
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
 Plate: LLAM11407 row: d column: 16
 High quality sequence stop: 727.

FEATURES

source Location/Qualifiers
1. .727
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:516399"
/tissue_type="medulla"
/lab_host="DH10B"
/clone_lib="NIH_MGC_119"
/note="Organ: brain; Vector: pCMV-SPORT6; Site: 1: NotI;
Site 2: EcoRV (destroyed); RNA source normal medulla from
anonymous male age 27. Library is oligo-dT primed and
directionally cloned (EcoRV site is destroyed upon
cloning). Average insert size 1.3 kb, insert size range
0.9-3 kb. Library is normalized and enriched for
full-length clones and was constructed by C. Gruber
(Invitrogen). Research Genetics tracking code 013. Note:
this is a NIH MGC Library."
BASE COUNT 161 a 238 c 204 g 124 t
ORIGIN

Query Match 27.1%; Score 495.2; DB 12; Length 727;
Best Local Similarity 98.9%; Pred. No. 6.2e-99;
Matches 530; Conservative 0; Mismatches 3; Indels 3; Gaps 3;
QY 1234 GGGGTGCCCTGGAGGGCGGCG-TCAACTCCATCAAGAACAGCTTTCTGGGCTCACCCCG 1292
Db 1 GGGGTGCCCTGGAGGGCGGCGCTCACTCCATCAAGAACAGCTTTCTGGGCTCACCCCG 60
QY 1293 CTTCCACCGCCGGAAGTGCAG-GTTCCGACGCGGAGGAGATGTCCAACTGACACCCAG 1351
Db 61 CTTCCACCGCCGGAAGTGCAGAGTTCCGAGCGCGGAGGAGATGTCCAACTGACACCCAG 120
QY 1352 AGTCGTCCTCCAGAGCTGGGAGAGTCTCTGGTTCGGACTTCATCAGCTGGAGAGG 1411
Db 121 AGTCGTCCTCCAGAGCTGGGAGAGTCTCTGGTTCGGACTTCATCAGCTGGAGAGG 180
QY 1412 AGCAGCAGATCTTCGTGGTTCATCAAGACAAACCTCTGAGCTCCATCAAGGCTGACATCG 1471
Db 181 AGCAGCAGATCTTCGTGGTTCATCAAGACAAACCTCTGAGCTCCATCAAGGCTGACATCG 240
QY 1472 TGACGCTTCCTGTCGATTCACAGTCTCAGCCACAGCGTCATCTCCCAAACGAGCTTCC 1531
Db 241 TGACGCTTCCTGTCGATTCACAGTCTCAGCCACAGCGTCATCTCCCAAACGAGCTTCC 300
QY 1532 GGGCCGAGTACAGGCCACAGGGGGCCAGCCGTTTCAGAGAGCCGTCAGTTCACAG 1591
Db 301 GGGCCGAGTACAGGCCACAGGGGGCCAGCCGTTTCAGAGAGCCGTCAGTTCACAG 360
QY 1592 TTGATATCAGCTACACCG-AGGGTGGGAGGCGCAGAGGAGAACGGCATCTACTCCGTC 1650
Db 361 TTGATATCAGCTACACCGTAGGTGGGAGGCGCAGAGGAGAACGGCATCTACTCCGTC 420
QY 1651 ACCTTACCTCTCTCAGGCCCCAGCGCTCGCTTCAAGAGGGTGGTGGAGACATCCAG 1710
Db 421 ACCTTACCTCTCTCAGGCCCCAGCGCTCGCTTCAAGAGGGTGGTGGAGACATCCAG 480
QY 1711 GCCCAGCTGTGAGCACACAGCCCGCTGCGCCCGAGCACTTGTTCAGACACCCAC 1766
Db 481 GCCCAGCTGTGAGCACACAGCCCGCTGCGCCCGAGCACTTGTTCAGACACCCAC 536

Search completed: November 27, 2003, 11:05:17
Job time : 3469.16 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 27, 2003, 09:04:30 ; Search time 1060.75 Seconds
(without alignments)
5674.682 Million cell updates/sec

Title: US-10-054-579-3
Perfect score: 1827
Sequence: 1 atgaagtgaggcgaggagat.....gaattatcccgaaagtaa 1827

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 2190069 seqs, 1647345023 residues
Total number of hits satisfying chosen parameters: 4380138

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications NA:*

- 1: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq:*
- 2: /cgn2_6/ptodata/2/pubpna/PCT_NEW_PUB.seq:*
- 3: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq:*
- 4: /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq:*
- 5: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq:*
- 6: /cgn2_6/ptodata/2/pubpna/PCTUS_PUBCOMB.seq:*
- 7: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq:*
- 8: /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq:*
- 9: /cgn2_6/ptodata/2/pubpna/US09A_PUBCOMB.seq:*
- 10: /cgn2_6/ptodata/2/pubpna/US09B_PUBCOMB.seq:*
- 11: /cgn2_6/ptodata/2/pubpna/US09C_PUBCOMB.seq:*
- 12: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq:*
- 13: /cgn2_6/ptodata/2/pubpna/US10A_PUBCOMB.seq:*
- 14: /cgn2_6/ptodata/2/pubpna/US10A_PUBCOMB.seq:*
- 15: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq:*
- 16: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq:*
- 17: /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1827	100.0	1827	13	US-10-054-579-3
2	1827	100.0	2007	13	US-10-054-579-1
3	1827	100.0	2647	12	US-10-288-798-42
4	1825.4	99.9	2908	14	US-10-195-072-1
5	1825.4	99.9	2908	14	US-10-195-071-1
6	1813	99.2	3364	14	US-10-195-072-3
7	1813	99.2	3364	14	US-10-195-071-3
8	1761.2	96.4	2025	14	US-10-283-247-1
9	1759.6	96.3	2025	10	US-09-842-582-3
10	1759.6	96.3	2219	10	US-09-842-582-1
11	1685.2	92.2	1911	14	US-10-283-247-4
12	851.4	46.6	2112	12	US-10-116-326-5
13	851.4	46.6	2337	12	US-10-116-326-1
14	851.4	46.6	2980	12	US-10-354-358-91
15	777.2	42.5	2289	12	US-10-116-326-3
16	594	32.5	2173	12	US-10-102-558-1

17	594	32.5	2173	14	US-10-102-558-1	Sequence 1, Appli
18	291.6	16.0	512	10	US-09-960-253-87	Sequence 87, Appli
19	232.4	12.7	1594	14	US-10-284-060-1	Sequence 1, Appli
20	232.4	12.7	1594	14	US-10-284-060-3	Sequence 3, Appli
21	225	12.3	2085	12	US-10-161-565-20	Sequence 20, Appli
22	225	12.3	2278	12	US-10-161-565-21	Sequence 21, Appli
23	225	12.3	3226	12	US-10-161-565-23	Sequence 23, Appli
24	225	12.3	3270	10	US-09-835-081-1	Sequence 1, Appli
25	225	12.3	3609	12	US-10-161-565-19	Sequence 19, Appli
26	225	12.3	4917	12	US-10-161-565-22	Sequence 22, Appli
27	222.8	12.2	2352	9	US-09-815-915-3	Sequence 3, Appli
28	222.8	12.2	2352	12	US-10-393-316-3	Sequence 3, Appli
29	222.8	12.2	2968	9	US-09-815-915-1	Sequence 1, Appli
30	232.8	12.2	2968	12	US-10-393-316-1	Sequence 1, Appli
31	214.2	11.7	26729	14	US-10-283-247-6	Sequence 6, Appli
32	214.2	11.7	70383	14	US-10-283-247-3	Sequence 3, Appli
33	193.8	10.6	2349	11	US-09-823-187-25	Sequence 25, Appli
34	182.4	10.5	3154	10	US-09-919-585-19	Sequence 19, Appli
35	187.6	10.3	1539	10	US-09-938-842A-2657	Sequence 2657, Ap
36	182.6	10.0	3609	9	US-09-799-875-6	Sequence 6, Appli
37	182.6	10.0	4381	12	US-10-288-798-31	Sequence 31, Appli
38	182.6	10.0	5983	9	US-09-799-875-4	Sequence 4, Appli
39	179	9.8	2086	12	US-10-161-565-4	Sequence 4, Appli
40	179	9.8	2112	10	US-09-919-585-4	Sequence 4, Appli
41	179	9.8	2112	12	US-10-161-565-7	Sequence 7, Appli
42	179	9.8	2222	10	US-09-919-585-7	Sequence 7, Appli
43	179	9.8	2248	12	US-10-161-565-5	Sequence 5, Appli
44	179	9.8	2701	12	US-10-161-565-6	Sequence 6, Appli
45	179	9.8	2784	12	US-10-161-565-2	Sequence 2, Appli

ALIGNMENTS

RESULT 1

US-10-054-579-3
; Sequence 3, Application US/10054579
; Publication No. US20020137913A1
; GENERAL INFORMATION:
; APPLICANT: Turner, C. Alexander Jr.
; APPLICANT: Mathuz, Brian
; TITLE OF INVENTION: No. US20020137913A1 Human Kinases and Polynucleotides Encoding t
; FILE REFERENCE: LEX-0300-USA
; CURRENT APPLICATION NUMBER: US/10/054,579
; CURRENT FILING DATE: 2002-01-22
; PRIOR APPLICATION NUMBER: US 60/263,378
; PRIOR FILING DATE: 2001-01-23
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 1827
; TYPE: DNA
; ORGANISM: homo sapiens
US-10-054-579-3

Query Match	100.0%;	Score 1827;	DB 13;	Length 1827;
Best Local Similarity	100.0%;	Pred. No. 0;		
Matches 1827;	Conservative	0;	Mismatches	0;
			Indels	0;
			Gaps	0;
Qy	1	ATGAAGTGGAGCGGAGATCGCGATCCTGAAGCTATTGAGCACCCCGCTCTAAAG	60	
Db	1	ATGAAGTGGAGCGGAGATCGCGATCCTGAAGCTATTGAGCACCCCGCTCTAAAG	60	
Qy	61	CTGCACGACGCTTTATGAAAACAAAATATTTGCTGCTGTAGAACACGCTCAGGT	120	
Db	61	CTGCACGACGCTTTATGAAAACAAAATATTTGCTGCTGTAGAACACGCTCAGGT	120	
Qy	121	GGTGAGCTCTTCGACTACCTGGTGAAGAGGGAGGCTGACGCCCTAAGAGGCTCGGAAG	180	
Db	121	GGTGAGCTCTTCGACTACCTGGTGAAGAGGGAGGCTGACGCCCTAAGAGGCTCGGAAG	180	
Qy	181	TTCTTCGGGAGATCATCTCTGCGCTGAGCTTTCGCACAGCCACTCATATGCGCAGG	240	
Db	181	TTCTTCGGGAGATCATCTCTGCGCTGAGCTTTCGCACAGCCACTCATATGCGCAGG	240	

Db 181 TTCTTCGGCAGATCATCTCTGCGCTGGAATCTTCTGCACAGCCACTCCATATGCCACAGG 240
Qy 241 GATCTGAACCTGAAACCTCTCTGCGAGAGAAACAACTCCGATCGCAGCTTT 300
Db 241 GATCTGAACCTGAAACCTCTCTGCGAGAGAAACAACTCCGATCGCAGCTTT 300
Qy 301 GGCATGGCGTCCCTGCGAGGTTGGCAGACGCTGTTGGAGACCAAGCTGTGGGTCCCCCAC 360
Db 301 GGCATGGCGTCCCTGCGAGGTTGGCAGACGCTGTTGGAGACCAAGCTGTGGGTCCCCCAC 360
Qy 361 TACGCTGCCCGAGGTGATTCGGGGGAGAGATGACGGCGGGAAGGGCGAGCTGTGG 420
Db 361 TACGCTGCCCGAGGTGATTCGGGGGAGAGATGACGGCGGGAAGGGCGAGCTGTGG 420
Qy 421 AGCTGGCGGTCATCTCTGCTGCTGCTGGGGCTCTGCGCTTCGACGATGACAAC 480
Db 421 AGCTGGCGGTCATCTCTGCTGCTGCTGGGGCTCTGCGCTTCGACGATGACAAC 480
Qy 481 TTGCGACAGCTGCTGGAGAGGTGAAGCGGGCGTGTTCACATGCGCCACTTTATCCCG 540
Db 481 TTGCGACAGCTGCTGGAGAGGTGAAGCGGGCGTGTTCACATGCGCCACTTTATCCCG 540
Qy 541 CCGACTGCCAGAGTCTCTACGGGGCATGATCGAGTGGACGGCGGCGCTCAAG 600
Db 541 CCGACTGCCAGAGTCTCTACGGGGCATGATCGAGTGGACGGCGGCGCTCAAG 600
Qy 601 CTAGAGCAATTCAGAAACATATGTTATAGGGGCAAGATGAGCCCGAACCAGAG 660
Db 601 CTAGAGCAATTCAGAAACATATGTTATAGGGGCAAGATGAGCCCGAACCAGAG 660
Qy 661 CAGCCCATTCCTGCAAGGTGAGATCGCTGCTGCTGCGCCAGCTTGGAGGACATCGACCC 720
Db 661 CAGCCCATTCCTGCAAGGTGAGATCGCTGCTGCTGCGCCAGCTTGGAGGACATCGACCC 720
Qy 721 GACGTCTGGACAGATGATCTGCTGGGCTGCTTCCGAGACCGCAACAGCTGCTGAG 780
Db 721 GACGTCTGGACAGATGATCTGCTGGGCTGCTTCCGAGACCGCAACAGCTGCTGAG 780
Qy 781 GACCTGCTGCTCGAGGAGAGAACAGAGAGATGATTACTTCTCTCTCGACCGG 840
Db 781 GACCTGCTGCTCGAGGAGAGAACAGAGAGATGATTACTTCTCTCTCGACCGG 840
Qy 841 AAAGAAAGGTACCCGAGCAGAGATGAGACCTGCCCCCGGAAACGAGATGACCT 900
Db 841 AAAGAAAGGTACCCGAGCAGAGATGAGACCTGCCCCCGGAAACGAGATGACCT 900
Qy 901 CCCCAGAGCGTGTGACTCCCGATGCTGACCGGCAAGCGCGGCGGCGGCGGCG 960
Db 901 CCCCAGAGCGTGTGACTCCCGATGCTGACCGGCAAGCGCGGCGGCGGCGGCG 960
Qy 961 AAATCCATGGAGGTGCTCAGCGTGACGACGGCGGCTCCCGGTCCTGCGCGGCGG 1020
Db 961 AAATCCATGGAGGTGCTCAGCGTGACGACGGCGGCTCCCGGTCCTGCGCGGCGG 1020
Qy 1021 ATTGAGATGGCCACAGCGGCGAGAGTCTCGGTTCATCAGGGTGCCTCTCAGGCGCTT 1080
Db 1021 ATTGAGATGGCCACAGCGGCGAGAGTCTCGGTTCATCAGGGTGCCTCTCAGGCGCTT 1080
Qy 1081 TCCACAGCCCACTCAGACGCCCCGGGTGACCCCTCACCCCTCACAAAGGGGCGATCCC 1140
Db 1081 TCCACAGCCCACTCAGACGCCCCGGGTGACCCCTCACCCCTCACAAAGGGGCGATCCC 1140
Qy 1141 CTCCACACCCCAAGGGGACACTGTTCACACGCCCAAGAGAGAGCCGGCTGGCAGCGCC 1200
Db 1141 CTCCACACCCCAAGGGGACACTGTTCACACGCCCAAGAGAGAGCCGGCTGGCAGCGCC 1200
Qy 1201 AACCCCAAGCCCGCTCAGCGCCAGAGTCTCGGTTCAGAGGGGTGCTTGGAGGGCGGCTCAAC 1260
Db 1201 AACCCCAAGCCCGCTCAGCGCCAGAGTCTCGGTTCAGAGGGGTGCTTGGAGGGCGGCTCAAC 1260
Qy 1261 TCCATCAAGAACAGCTTTCTGGGCTCACCCCGCTTCCACCGCGGAAACTGCAAGTTCCG 1320
Db 1261 TCCATCAAGAACAGCTTTCTGGGCTCACCCCGCTTCCACCGCGGAAACTGCAAGTTCCG 1320

Qy 1321 AGCCCGAGAGATGTCCAACTGTACACAGAGTGTCTCCAGAGCTGGCGGAAGTCC 1380
Db 1321 AGCCCGAGAGATGTCCAACTGTACACAGAGTGTCTCCAGAGCTGGCGGAAGTCC 1380
Qy 1381 TGGTTGGGAACTTCATCAGCTTGAAGAGGAGAGAGATCTTCTGTGTCTATCAAGAC 1440
Db 1381 TGGTTGGGAACTTCATCAGCTTGAAGAGGAGAGAGATCTTCTGTGTCTATCAAGAC 1440
Qy 1441 AAACCTCTCAGCTTCCATCAAGCTGACGCTGACGCTTCTGTGCTATCCAGTCTC 1500
Db 1441 AAACCTCTCAGCTTCCATCAAGCTGACGCTTCTGTGCTATCCAGTCTC 1500
Qy 1501 AGCCACAGAGCTCATCTCCCAAGAGAGCTTCCGGGCGGAGTACAAGGCCACGGGGGGCCA 1560
Db 1501 AGCCACAGAGCTCATCTCCCAAGAGAGCTTCCGGGCGGAGTACAAGGCCACGGGGGGCCA 1560
Qy 1561 GCGTGTTCAGAGAGCGGTCAAGTTCAGGTTCCAGTTGATATCACCCTACACGAGGGTGGGAG 1620
Db 1561 GCGTGTTCAGAGAGCGGTCAAGTTCAGGTTCCAGTTGATATCACCCTACACGAGGGTGGGAG 1620
Qy 1621 GCGCAGAGAGAGAGCGCATCTACTCCGTCACCTTCACTCTCAGCCCTCTCAGCCCGGCGT 1680
Db 1621 GCGCAGAGAGAGAGCGCATCTACTCCGTCACCTTCACTCTCAGCCCTCTCAGCCCGGCGT 1680
Qy 1681 CGCTTCAAGAGAGGTGGAGACCATCCAGGCCAGCTGCTGAGCACACACGACCCGCT 1740
Db 1681 CGCTTCAAGAGAGGTGGAGACCATCCAGGCCAGCTGCTGAGCACACACGACCCGCT 1740
Qy 1741 GCGGCCGACACTTGTACAGACACCATTAAGTATGTAAGAAATGATGAGGGGCGGCTTCC 1800
Db 1741 GCGGCCGACACTTGTACAGACACCATTAAGTATGTAAGAAATGATGAGGGGCGGCTTCC 1800
Qy 1801 AAATGTGGAATATCCCGAAAGTTAA 1827
Db 1801 AAATGTGGAATATCCCGAAAGTTAA 1827

RESULT 2
US-10-054-579-1
; Sequence 1, Application US/10054579
; Publication No. US20020137913A1
; GENERAL INFORMATION:
; APPLICANT: Turner, C. Alexander Jr.
; TITLE OF INVENTION: No. US20020137913A1 Human Kinases and Polynucleotides Encoding t
; FILE REFERENCE: LEX-0300-USA
; CURRENT APPLICATION NUMBER: US/10/054,579
; PRIOR FILING DATE: 2002-01-22
; PRIOR APPLICATION NUMBER: US 60/263,378
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: Fast-Seq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 2007
; TYPE: DNA
; ORGANISM: homo sapiens
US-10-054-579-1

Query Match 100.0%; Score 1827; DB 13; Length 2007;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1827; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ATGAAGTGGAGGGGAGATCGGATCTCTGAAGCTCATTAAGCTCATTAAGCAACCCCGCTCTAAAG 60
Db 181 ATGAAGTGGAGGGGAGATCGGATCTCTGAAGCTCATTAAGCAACCCCGCTCTAAAG 240
Qy 61 CTGACAGCTTTATGAAAACAAAATATTTGATCTGTGTGTAGAACACGTTGCTAGGT 120
Db 241 CTGACAGCTTTATGAAAACAAAATATTTGATCTGTGTGTAGAACACGTTGCTAGGT 300
Qy 121 GGTGAGCTTCTCGACTACCTGCTGAGAGGGGAGGCTGACGCTTAAGGAGGCTCGGAG 180
Db 121 GGTGAGCTTCTCGACTACCTGCTGAGAGGGGAGGCTGACGCTTAAGGAGGCTCGGAG 180

301	GGTGAGCTCTTGACATCACTGTGTGAAGAGGGAGGCTGACGCTTAAGGAGGCTCGGAAG	360
181	TTCTTCCGGCAGATCATCTCTGGCGTGGACTTCTGCCACAGCCACTCCATATGCCACAGG	240
	TTCTTCCGGCAGATCATCTCTGGCTGGACTTCTGCCACAGCCACTCCATATGCCACAGG	420
241	GATCTGAACCTGAAAACCTCTCTGTCTGGACGAGAGAA	300
421	GATCTGAACCTGAAAACCTCTCTGTCTGGACGAGAGAA	360
	GATCTGAACCTGAAAACCTCTCTGTCTGGACGAGAGAA	480
301	GGCATGCGCTCCCTGCAGAGTTGCGCAGACCTGTTGGAGACCAAGCTGTGGTCCCCCCAC	360
481	GGCATGCGCTCCCTGCAGAGTTGCGCAGACCTGTTGGAGACCAAGCTGTGGTCCCCCCAC	540
	TACGCTTCCGCGGAGGTGATCCGCGGGGAGAAGTATGACGCGCGGAGGCGGACGTGTGG	420
541	TACGCTTCCGCGGAGGTGATCCGCGGGGAGAAGTATGACGCGCGGAGGCGGACGTGTGG	600
421	AGCTGGCGGTCACTCTGTTTTCGCTTGTCTGGTGGGGCTCTGCGCTTTCGACGATGACAAC	480
	AGCTGGCGGTCACTCTGTTTTCGCTTGTCTGGTGGGGCTCTGCGCTTTCGACGATGACAAC	660
481	TTCCGACAGCTGCTGGAGAGGTGACGCGGGGCGTGTTCACATCGCGCACTTTATCCCG	540
661	TTCCGACAGCTGCTGGAGAGGTGACGCGGGGCGTGTTCACATCGCGCACTTTATCCCG	720
	CCGCACTGCCAGAGTCTGCTACCGGGGCAATGATCGAGTGGACGCCGACGCCCTCAAG	600
721	CCGCACTGCCAGAGTCTGCTACCGGGGCAATGATCGAGTGGACGCCGACGCCCTCAAG	780
	CTAGAGCACTTCAGAAACACATATGGTATATAGGGGGCAAGAAATGAGCCGCAACAGAG	660
781	CTAGAGCACTTCAGAAACACATATGGTATATAGGGGGCAAGAAATGAGCCGCAACAGAG	840
661	CAGCCATTCTTCGCAAGTGCAGATCCGCTCCCTGCCAGCTCGAGGACATCGACCCC	720
	CAGCCATTCTTCGCAAGTGCAGATCCGCTCCCTGCCAGCTCGAGGACATCGACCCC	900
721	GAGCTGTGCACAGCATGCATCACTGGGCTGCTTCCGAGACCGCAACAGCTGCTGCAG	780
901	GAGCTGTGCACAGCATGCATCACTGGGCTGCTTCCGAGACCGCAACAGCTGCTGCAG	960
	GACCTGTCTCCGAGGAGGAGAAC	840
961	GACCTGTCTCCGAGGAGGAGAAC	1020
841	AAAGAAAGGTACCCGAGCCAGGAGATGAGGACTCGCCCGCGGACGAGATAGACCTT	900
	AAAGAAAGGTACCCGAGCCAGGAGATGAGGACTCGCCCGCGGACGAGATAGACCTT	1080
901	CCCCGGAAGCGTGTGGACTCCCGATGCTGAAACCGGACGCGCAAGCGGCGGCGCAGAACGC	960
1081	CCCCGGAAGCGTGTGGACTCCCGATGCTGAAACCGGACGCGCAAGCGGCGGCGCAGAACGC	1140
	AAATCCATGAGGTGCTCAGCGTGACGAGCGGGGCTCCCGGTGCTCGCGCGCGGGCC	1020
1141	AAATCCATGAGGTGCTCAGCGTGACGAGCGGGGCTCCCGGTGCTCGCGCGCGGGCC	1200
1021	ATTGAGATGCCCCAGCAGCGGCGAGAGTCTCGGTTCATGACGGGTGCTCTCTCAGGCTTT	1080
	ATTGAGATGCCCCAGCAGCGGCGAGAGTCTCGGTTCATGACGGGTGCTCTCTCAGGCTTT	1260
1081	TCCACAGCCCACTCAGCAGCCCCCGGTGACCCCTACCCCTCACCAAGGGGCACTCCC	1140
1261	TCCACAGCCCACTCAGCAGCCCCCGGTGACCCCTACCCCTCACCAAGGGGCACTCCC	1320
	CTCCCCACCCCAAGGGGCACTCTGTCACACGCGCAAGGAGAGCCCGGTGCGCAGCCCC	1200
1321	CTCCCCACCCCAAGGGGCACTCTGTCACACGCGCAAGGAGAGCCCGGTGCGCAGCCCC	1380
1201	AAACCCACGCCCCCGTCCAGCCCCAGCGTCCGAGGGGTGCTCTGAGGGGCGCGGCTCAAC	1260
1381	AAACCCACGCCCCCGTCCAGCCCCAGCGTCCGAGGGGTGCTCTGAGGGGCGCGGCTCAAC	1440

RESULT 3

US-10-288-798-42

US-10-288-798-42
: Sequence 42 Application IIS/10288798

; sequence 42, Application US/102
: Publication No. US20030207299A1; Publication No: US20
: GENERAL INFORMATION:

/ GENERAL INFORMATION:
 / APPLICANT: BANDMAN, Olga; NGUYEN, Damnie B;
 / APPLICANT: WALIA, Narinder K.; HAPALITA, April J.A.;
 / APPLICANT: YAO, Monique G.; GANDHI, Ameena R.;
 / APPLICANT: GURURAJAN, Rajagopal.; DING, Li;
 / APPLICANT: PATTERSON, Chandra.; YUE, Henry;
 / APPLICANT: BAUGHN, Mariah R.; TRIBOULEY, Catherine M.;
 / APPLICANT: THORNTON, Michael.; ELLIOTT, Vicki S.;
 / APPLICANT: LU, Yan.; ISON, Craig H.;
 / APPLICANT: AU-YOUNG, Janice.; TANG, Y. Tom;
 / APPLICANT: ARIZWALI, Yalga.; BURRILL, John D.;
 / APPLICANT: MACUS, Gregory A.; ZINGLER, Kurt A.;
 / APPLICANT: LU, Dying Aina M.; LAL, Preeti G.;
 / APPLICANT: RAMKUMAR, Jayalaxmi.; WARREN, Bridget A.;
 / APPLICANT: KEARNEY, Liam.; POLICKY, Jennifer L.;
 / APPLICANT: THANGAVELU, Kavitcha.; BURFORD, Neil
 / TITLE OF INVENTION: HUMAN KINASES
 / FILE REFERENCE: PI-0209 USA
 / CURRENT APPLICATION NUMBER: US/10/288,798
 / CURRENT FILING DATE: 2002-11-01
 / PRIOR APPLICATION NUMBER: PCT/US01/27219
 / PRIOR FILING DATE: 2001-08-31
 / PRIOR APPLICATION NUMBER: US 60/240,542
 / PRIOR FILING DATE: 2000-10-13
 / PRIOR APPLICATION NUMBER: US 60/238,389
 / PRIOR FILING DATE: 2000-10-06
 / PRIOR APPLICATION NUMBER: US 60/236,499

;; PRIOR FILING DATE: 2000-09-29
;; PRIOR APPLICATION NUMBER: US 60/234,902
;; PRIOR FILING DATE: 2000-09-22
;; PRIOR APPLICATION NUMBER: US 60/232,654
;; PRIOR FILING DATE: 2000-09-14
;; PRIOR APPLICATION NUMBER: US 60/231,357
;; PRIOR FILING DATE: 2000-09-08
;; PRIOR APPLICATION NUMBER: US 60/229,873
;; PRIOR FILING DATE: 2000-08-31
;; NUMBER OF SEQ ID NOS: 48
;; SOFTWARE: PERL Program
;; SEQ ID NO 42
;; LENGTH: 2647
;; TYPE: DNA
;; ORGANISM: Homo sapiens
;; FEATURE:
;; NAME/KEY: misc.feature
;; OTHER INFORMATION: Incyte ID No. US20030207299A1 4022651CB1
US-10-288-798-42

Query Match 100.0%; Score 1827; DB 12; Length 2647;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1827; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ATGAAGGTGGAGCGGGAGATCGCGATCCTGAAGCTCATTTAGACACCCGCCACGTCCTAAAG 60
Db 169 ATGAAGGTGGAGCGGGAGATCGCGATCCTGAAGCTCATTTAGACACCCGCCACGTCCTAAAG 228

Qy 61 CTGACGACGTTTATGAAAAAATAATTTGTAACCTGTGTGTAGAACACGTTGTACGGT 120
Db 229 CTGACGACGTTTATGAAAAAATAATTTGTAACCTGTGTGTAGAACACGTTGTACGGT 288

Qy 121 GGTGAGCTCTTCGACTACTCTGCTGAAGAGGGAGGCTGACGCCCTAAGGAGGCTCGGAAG 180
Db 289 GGTGAGCTCTTCGACTACTCTGCTGAAGAGGGAGGCTGACGCCCTAAGGAGGCTCGGAAG 348

Qy 181 TTCTTCGGCGAGATCATCTCTGCGCTGAGACTTTCTGCCACAGCCACTCCATATGCCACAGG 240
Db 349 TTCTTCGGCGAGATCATCTCTGCGCTGAGACTTTCTGCCACAGCCACTCCATATGCCACAGG 408

Qy 241 GATCTGAACCTGAAACCTCTCTGCTGGAAGAGAAACAAACATCCGATCGAGACTTT 300
Db 409 GATCTGAACCTGAAACCTCTCTGCTGGAAGAGAAACAAACATCCGATCGAGACTTT 468

Qy 301 GGCAATGGCGTCCCTGCAAGTGGGACAGCTGTTGGAGACCACTGTGGGTCCGCCAC 360
Db 469 GGCAATGGCGTCCCTGCAAGTGGGACAGCTGTTGGAGACCACTGTGGGTCCGCCAC 528

Qy 361 TACGCTGCCCCGAGGTGATCCGGGGGAGAGATATGACGCGGAAAGCGGAGCGTGTGG 420
Db 529 TACGCTGCCCCGAGGTGATCCGGGGGAGAGATATGACGCGGAAAGCGGAGCGTGTGG 588

Qy 421 AGCTGGCGGTTCATCTGCTTGGCTTGGTGGGGGCTGTCGCCCTTCGACGATGACAC 480
Db 589 AGCTGGCGGTTCATCTGCTTGGCTTGGTGGGGGCTGTCGCCCTTCGACGATGACAC 648

Qy 481 TTGGCGACAGCTGCTGGAGAAAGTGAAGCGGGCGGTGTTCCACATGCCGCATTTATCCCG 540
Db 649 TTGGCGACAGCTGCTGGAGAAAGTGAAGCGGGCGGTGTTCCACATGCCGCATTTATCCCG 708

Qy 541 CCCGACTGCCAGAGTCTGCTACGCGGGCATATCGAGGTGGAGCGCGCACGCCCTCACG 600
Db 709 CCCGACTGCCAGAGTCTGCTACGCGGGCATATCGAGGTGGAGCGCGCACGCCCTCACG 768

Qy 601 CTAGAGCATTTCAGAAACATATGTTATATAGGGGCGAGAAATGAGCCCGAACCCAGAG 660
Db 769 CTAGAGCATTTCAGAAACATATGTTATATAGGGGCGAGAAATGAGCCCGAACCCAGAG 828

Qy 661 CAGCCCATTCCTCGCAAGGTGCAGATCCGCTCGCTGCGTGGAGGACATCGACCCC 720
Db 829 CAGCCCATTCCTCGCAAGGTGCAGATCCGCTCGCTGCGTGGAGGACATCGACCCC 888

Qy 721 GACGTGCTGGAAGCATGCACTCACTGGGCTGCTTTCGAGAGCCGCAACAAAGCTGCTGAG 780

Db 889 GACGTGCTGGAAGCATGCACTCACTGGGCTGCTTCCGAGACCGCAACAGCTGCTGCAG 948
Qy 781 GACCTGCTGTCGAGGAGAGAACAGGAGAGATGATTACTTCTCTCTCTGAGACCGG 840
Db 949 GACCTGCTGTCGAGGAGAGAACAGGAGAGATGATTACTTCTCTCTCTGAGACCGG 1008
Qy 841 AAAGAAAGGTACCAGAGCCAGGAGGATGAGGACTCCGCCCGGAAACAGATAGACCT 900
Db 1009 AAAGAAAGGTACCCGAGCCAGGAGGATGAGGACTCCGCCCGGAAACAGATAGACCT 1068
Qy 901 CCCCAGAGCGTGTGGAATCCCGATGCTGAACCGGACCGGACCGGCGGCGGCGGCG 960
Db 1069 CCCCAGAGCGTGTGGAATCCCGATGCTGAACCGGACCGGAGCGGCGGCGGCGGCG 1128
Qy 961 AATTCATGAGGTGCTCAGCGTGACGAGCGGCGGCTCCCGGTGCTTGGCGGCGGCG 1020
Db 1129 AATTCATGAGGTGCTCAGCGTGACGAGCGGCGGCTCCCGGTGCTTGGCGGCGGCG 1188
Qy 1021 ATTGAGATGGCCAGCAGCAGCGGAGAGGTCTCGGTCCATCAGCGGTGCTCTCAGGCGCTT 1080
Db 1189 ATTGAGATGGCCAGCAGCAGCGGAGAGGTCTCGGTCCATCAGCGGTGCTCTCAGGCGCTT 1248
Qy 1081 TCACACAGCCCACTCAGCAGCCCGCGGTGAACCCCTCAACCCCTCAACAGAGGCGAGTCC 1140
Db 1249 TCACACAGCCCACTCAGCAGCCCGCGGTGAACCCCTCAACCCCTCAACAGAGGCGAGTCC 1308
Qy 1141 CTCGCCACCCCAAGGGGACCTGTGCACAGCCCAAGAGAGAGCGGCTGCGACGCGCC 1200
Db 1309 CTCGCCACCCCAAGGGGACCTGTGCACAGCCCAAGAGAGAGCGGCTGCGACGCGCC 1368
Qy 1201 AACCCCAAGCGGCTGACGCGCCAGCGTCGAGAGGCGTCCCTGAGAGGCGCGGCTCAAC 1260
Db 1369 AACCCCAAGCGGCTGACGCGCCAGCGTCGAGAGGCGTCCCTGAGAGGCGCGGCTCAAC 1428
Qy 1261 TCATCAAGAAACAGCTTTCTGGGCTCAACCCCGTTCACCGCGGAAACCTGCAAGTTCCG 1320
Db 1429 TCATCAAGAAACAGCTTTCTGGGCTCAACCCCGTTCACCGCGGAAACCTGCAAGTTCCG 1488
Qy 1321 ACCCGGAGAGATGTCACACCTGACAGAGTCTGCCAGAGCTGCGCAAGAGAGTCC 1380
Db 1489 ACCCGGAGAGATGTCACACCTGACAGAGTCTGCCAGAGCTGCGCAAGAGAGTCC 1548
Qy 1381 TGCTTTGGGAACTTCATCAGCCTGGAGAGGAGGAGCAGATCTTCGTGGTTCATCAAGAC 1440
Db 1549 TGCTTTGGGAACTTCATCAGCCTGGAGAGGAGGAGCAGATCTTCGTGGTTCATCAAGAC 1608
Qy 1441 AAACCTCTGAGTCTCATCAAGGCTGACGCGCTTCCTGTCGATTCCTCCAGTCTC 1500
Db 1609 AAACCTCTGAGTCTCATCAAGGCTGACGCGCTTCCTGTCGATTCCTCCAGTCTC 1668
Qy 1501 AGCCACAGCGTCTATCTCCAAACGAGCTTCCGGGCGGAGTACAAAGGCCACGGGGGGCCA 1560
Db 1669 AGCCACAGCGTCTATCTCCAAACGAGCTTCCGGGCGGAGTACAAAGGCCACGGGGGGCCA 1728
Qy 1561 GCGGTGTTCCAGAAAGCGGTTCAGTTTCAGGTTGATATCACCTACACGAGGTTGGGAG 1620
Db 1729 GCGGTGTTCCAGAAAGCGGTTCAGTTTCAGGTTGATATCACCTACACGAGGTTGGGAG 1788
Qy 1621 GCGCAGAGAGAAAGGCACTCTACTCGGTCACTTCACCTGCTCTCAGGCGCCCGCGCT 1680
Db 1789 GCGCAGAGAGAAAGGCACTCTACTCGGTCACTTCACCTGCTCTCAGGCGCCCGCGCT 1848
Qy 1681 CGCTTCAAGAGGTTGGAGACCATCCAGGCGGAGTCCAGGTCGTGAGCACAACAGCCCGCT 1740
Db 1849 CGCTTCAAGAGGTTGGAGACCATCCAGGCGGAGTCCAGGTCGTGAGCACAACAGCCCGCT 1908
Qy 1741 GCGGCGGCGAGCATTTGTGAGACACCACTAACTGATGGAATGATGAGGCGGCGGCTTCC 1800
Db 1909 GCGGCGGCGAGCATTTGTGAGACACCACTAACTGATGGAATGATGAGGCGGCGGCTTCC 1968
Qy 1801 AATGTGGAATTTATCCCGAAAAAGTTAA 1827

QY 1741 GCGGCCAGACCTTGTGACACACCACTAACTGTATGAAATGATGACGGGCGGCTTTCC 1800
Db 2026 GCGGCCAGACCTTGTGACACACCACTAACTGTATGAAATGATGACGGGCGGCTTTCC 2085
QY 1801 AAATGTGGAATTATCCCGAAAGTTAA 1827
Db 2086 AAATGTGGAATTATCCCGAAAGTTAA 2112

RESULT 5
US-10-195-071-1
; Sequence 1, Application US/10195071
; Publication No. US20030096271A1
; GENERAL INFORMATION:
; APPLICANT: Origene Technologies
; TITLE OF INVENTION: Full-Length Serine Protein Kinase in Brain and Pancreas
; FILE REFERENCE: 160 101 C1
; CURRENT APPLICATION NUMBER: US/10/195,071
; CURRENT FILING DATE: 2002-07-15
; PRIOR APPLICATION NUMBER: US 09/930,181
; PRIOR FILING DATE: 2001-08-16
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1
; LENGTH: 2908
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (106)..(2112)
; OTHER INFORMATION:
US-10-195-071-1

Query Match 99.9%; Score 1825.4; DB 14; Length 2908;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1826; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 ATGAAGTGGAGCGGAGATCCGATCTCTGAAGCTCAATGAGACACCCCGCTCTAAAG 60
Db 286 ATGAAGTGGAGCGGAGATCCGATCTCTGAAGCTCAATGAGACACCCCGCTCTAAAG 345
QY 61 CTGACGACGTTTATGAAACAAATAATTTGTACCTGTGTGTAGAACACGTTCTAGGT 120
Db 346 CTGACGACGTTTATGAAACAAATAATTTGTACCTGTGTGTAGAACACGTTCTAGGT 405
QY 121 GGTGAGCTCTTCGACTCTGTTGAAGAGGGAGGCTGACGCTTAGAGGCTCGGAAG 180
Db 406 GGTGAGCTCTTCGACTCTGTTGAAGAGGGAGGCTGACGCTTAGAGGCTCGGAAG 465
QY 181 TTCTTCGCGCAGATCATCTCTGCGCTGGACTTTCTGCCACAGCCACTCCATATGCCACAGG 240
Db 466 TTCTTCGCGCAGATCATCTCTGCGCTGGACTTTCTGCCACAGCCACTCCATATGCCACAGG 525
QY 241 GATCTGAAACCTGAAACCTCTCTGTGGAAGAGAAACAAATCCGATCGCAGACTTT 300
Db 526 GATCTGAAACCTGAAACCTCTCTGTGGAAGAGAAACAAATCCGATCGCAGACTTT 585
QY 301 GGCATGGCGTCCCTGACGTTGGCAGACAGCTTTGAGAGACAGCTGTGGTCCCGCCAC 360
Db 586 GGCATGGCGTCCCTGACGTTGGCAGACAGCTTTGAGAGACAGCTGTGGTCCCGCCAC 645
QY 361 TACGCTTCCCGAGGTGATCCGGGGGAGAGTATGACGCGCGGAGCGGAGCGTGTGG 420
Db 646 TACGCTTCCCGAGGTGATCCGGGGGAGAGTATGACGCGCGGAGCGGAGCGTGTGG 705
QY 421 AGCTCGGGGTTCATCTGTTGCCCTTGTGTGGGGCTCTGCCCTTCGACGATGACAAAC 480
Db 706 AGCTCGGGGTTCATCTGTTGCCCTTGTGTGGGGGTCTGCCCTTCGACGATGACAAAC 765
QY 481 TTGCGACAGCTCTCGAGAGGTGAAGCGGGCGGTTCACATCCGCACTTTATCCCG 540
Db 766 TTGCGACAGCTCTCGAGAGGTGAAGCGGGCGGTTCACATCCGCACTTTATCCCG 825

QY 541 CCCGACTGCCAGTCTGTACGGGCGCATGATCGAGGTGGACGCGCCACGCGGCTCTCAG 600
Db 826 CCCGACTGCCAGTCTGTACGGGCGCATGATCGAGGTGGACGCGCCACGCGGCTCTCAG 885
QY 601 CTAGAGCACAATTCAGAAACA CATATGTTATATAGGGGGCAGAAATGAGCCCCGAAACAGAG 660
Db 886 CTAGAGCACAATTCAGAAACA CATATGTTATATAGGGGGCAGAAATGAGCCCCGAAACAGAG 945
QY 661 CAGCCCATTCCTCGCAAGGTGCAGATCCGCTCGCTGCGCCAGCTCGAGGACATCGACCCC 720
Db 946 CAGCCCATTCCTCGCAAGGTGCAGATCCGCTCGCTGCGCCAGCTCGAGGACATCGACCCC 1005
QY 721 GACGTGCTGGACAGCATGCACTCACTGGGCTGCTTCCGAGACCGCAACAAAGTGTGCGAG 780
Db 1006 GACGTGCTGGACAGCATGCACTCACTGGGCTGCTTCCGAGACCGCAACAAAGTGTGCGAG 1065
QY 781 GACCTGCTGTCGAGGAGGAGAACACGAGAGATGATTTACTTCTCTCTCTGAGACCGG 840
Db 1066 GACCTGCTGTCGAGGAGGAGAACACGAGGAGATGATTTACTTCTCTCTCTGAGACCGG 1125
QY 841 AAAGAAAGGTACCCGAGCCAGGAGATGAGGACCTGCCCCCGGAAACGAGATAGACCTT 900
Db 1126 AAAGAAAGGTACCCGAGCCAGGAGATGAGGACCTGCCCCCGGAAACGAGATAGACCTT 1185
QY 901 CCCCGAAGCGTGTGGACTTCCCGATGCTTGAACCGGACCGGCAAGCGGCGGCGAGAACGC 960
Db 1186 CCCCGAAGCGTGTGGACTTCCCGATGCTTGAACCGGACCGGCAAGCGGCGGCGAGAACGC 1245
QY 961 AATTCATCGAGGTGCTCAGCGTGAAGGAGCGGGCTCCCGGTGCTCGGCGGGGGCC 1020
Db 1246 AAGTCCATGGAGTGTCTCAGCGTGAAGGAGCGGGCTCCCGGTGCTCGGCGGGGGCC 1305
QY 1021 ATTGAGATGGCCAGCACGCGCCAGAGGTCTCGGTCCATCAGCGGTGCTCTCAGGCGCTT 1080
Db 1306 ATTGAGATGGCCAGCACGCGCCAGAGGTCTCGGTCCATCAGCGGTGCTCTCAGGCGCTT 1365
QY 1081 TCACACAGCCCACTCAGCAGCCCGCGGTGACCCCTCACCCCTCACCAAGGGGCGAGTCC 1140
Db 1366 TCACACAGCCCACTCAGCAGCCCGCGGTGACCCCTCACCCCTCACCAAGGGGCGAGTCC 1425
QY 1141 CTCGCCACCCCAAGGGGACACTGCTGCCACAGCCCAAGAGAGAGCCGCTGCGCAGCC 1200
Db 1426 CTCGCCACCCCAAGGGGACACTGCTGCCACAGCCCAAGAGAGAGCCGCTGCGCAGCC 1485
QY 1201 AACCCCAACGCGCCGCTCCAGCCCGAGCGGAGGCTGCTGAGAGGCGCGGCTCAAC 1260
Db 1486 AACCCCAACGCGCCCGCTCCAGCCCGAGCGGAGGCTGCTGAGAGGCGCGGCTCAAC 1545
QY 1261 TCATCAAGAACAGCTTTCTGGGCTCACCCCGTTCACCGCCGGAACCTGCAAGTTCCG 1320
Db 1546 TCATCAAGAACAGCTTTCTGGGCTCACCCCGTTCACCGCCGGAACCTGCAAGTTCCG 1605
QY 1321 AGCGCGGAGAGATGTCACACTGACACAGAGTCTGCTCCCGAGAGCTGGCGAGAGAGTCC 1380
Db 1606 AGCGCGGAGAGATGTCACACTGACACAGAGTCTGCTCCCGAGAGCTGGCGAGAGAGTCC 1665
QY 1381 TGGTTTGGGAACCTTCATCAGCTTGGAGAGAGGAGCAGATCTTCGTGTGTCATCAAGAG 1440
Db 1666 TGGTTTGGGAACCTTCATCAGCTTGGAGAGAGGAGCAGATCTTCGTGTGTCATCAAGAG 1725
QY 1441 AAACCTCTGAGCTCCATCAAGGCTGACATCGTGCAAGCTTCTGTGATCCAGTCTC 1500
Db 1726 AAACCTCTGAGCTCCATCAAGGCTGACATCGTGCAAGCTTCTGTGATCCAGTCTC 1785
QY 1501 AGCCACAGCTCTCTCCCAACAGAGCTTCCGGCGCGAGTACAAAGGCCACGGGGGGCCCA 1560
Db 1786 AGCCACAGCTCTCTCTCCCAACAGAGCTTCCGGCGCGAGTACAAAGGCCACGGGGGGCCCA 1845
QY 1561 GCGGTGTTCCAGAGCGGTCAGGTTCCAGGTTGATATACCTACACGAGGGGTGGGGAG 1620
Db 1846 GCGGTGTTCCAGAGCGGTCAGGTTCCAGGTTGATATACCTACACGAGGGGTGGGGAG 1905

Qy 1621 GCGAGAGGAGAAACGGCATCTTACCTCGTCACTTACCTGCTCTCAGGCCCCAGCCGT 1680
Db 1906 GCGCAGAGGAGAAACGGCATCTTACTCGTCACTTACCTGCTCTCAGGCCCCAGCCGT 1965
Qy 1681 CGCTTCAGAGGGTGGTGGAGACCATCAGGCCCGAGCTGCTGAGCAGACAGCCCGCT 1740
Db 1966 CGCTTCAGAGGGTGGTGGAGACCATCAGGCCCGAGCTGCTGAGCAGACAGCCCGCT 2025
Qy 1741 GCGGCCACAGCACTTGTTCAGACACCACTAACTGTATGGAATGATGACGCGGCGGCTTTCC 1800
Db 2026 GCGGCCACAGCACTTGTTCAGACACCACTAACTGTATGGAATGATGACGCGGCGGCTTTCC 2085
Qy 1801 AATGTGGAATATCCCGAAAGTTAA 1827
Db 2086 AATGTGGAATATCCCGAAAGTTAA 2112

RESULT 6
US-10-195-072-3
; Sequence 3, Application US/10195072
; Publication No. US20030092036A1
; GENERAL INFORMATION:
; APPLICANT: Origene Technologies
; TITLE OF INVENTION: Full-Length Serine Protein Kinase in Brain and Pancreas
; FILE REFERENCE: 160 101 C2
; CURRENT APPLICATION NUMBER: US/10/195,072
; PRIOR FILING DATE: 2002-07-15
; PRIOR APPLICATION NUMBER: US 09/930,181
; PRIOR FILING DATE: 2001-08-16
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: Patentin version 3.1
; SEQ ID NO 3
; LENGTH: 3364
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (482)..(2239)
; OTHER INFORMATION:
US-10-195-072-3

Query Match 99.2%; Score 1813; DB 14; Length 3364;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 1827; Conservative 0; Mismatches 0; Indels 4; Gaps 1;
Qy 1 ATGAAGGTGGAGCGGAGATCGCGATCCTGAAGCTCATTTGAGCACCCCGACGCTCTAAAG 60
Db 409 ATGAAGGTGGAGCGGAGATCGCGATCCTGAAGCTCATTTGAGCACCCCGACGCTCTAAAG 468
Qy 61 CTGACGACGCTTTATGAAAAAATAATTTT---GTACTGTGTGTAGAACACGCTGTC 116
Db 469 CTGACGACGCTTTATGAAAAAATAATTTTGTAGGTACTGTGTGTAGAACACGCTGTC 528
Qy 117 AGGTGGTGGAGCTTCGACTACCTGGTGAAGAGGGAGGCTGAGCGCTAAGGAGGCTCG 176
Db 529 AGGTGGTGGAGCTTCGACTACCTGGTGAAGAGGGAGGCTGAGCGCTAAGGAGGCTCG 588
Qy 177 GAAATTTCTTCGGGAGATCATCTCTGCGCTGGACTTTCTGCGACAGCCACTCCATATGCCA 236
Db 589 GAAATTTCTTCGGGAGATCATCTCTGCGCTGGACTTTCTGCGACAGCCACTCCATATGCCA 648
Qy 237 CAGGATCTGAAAACTTGAAAACTTCTGCTGGA CGAGAAAGAACATCCGATCGCAGA 296
Db 649 CAGGATCTGAAAACTTGAACCTCTCTGGA CGAGAAAGAACATCCGATCGCAGA 708
Qy 297 CTTTGGCATGGCTCCCTGCGAGTTGGCGACAGCTGTTTGAGACCACTGTGGTCCCC 356
Db 709 CTTTGGCATGGCTCCCTGCGAGTTGGCGACAGCTGTTTGAGACCACTGTGGTCCCC 768
Qy 357 CCATACCGCTGCCCGGAGGTGATCCGGGGGAGAGTATGACGCGCGGAAGGCGGAGCT 416
Db 769 CCATACCGCTGCCCGGAGGTGATCCGGGGGAGAGTATGACGCGCGGAAGGCGGAGCT 828

Qy 417 GTGAGCTGGGGCTCATCTGCTTGTGGTGGGGCTCTGCTCTCGACATGA 476
Db 829 GTGAGCTGGGGCTCATCTGCTTGTGGTGGGGCTCTGCTCTCGACATGA 888
Qy 477 CAACTTGGGACAGCTGCTGGAGAAAGTGAAGCGGGGCTGTTTCA CATATGCGGCACTTTAT 536
Db 889 CAACTTGGGACAGCTGCTGGAGAAAGTGAAGCGGGGCTGTTTCA CATATGCGGCACTTTAT 948
Qy 537 CCGCGCCGACTGCGACAGCTGCTAGCGGGCATGATCGAGGTGAGACGCGCGCCCT 596
Db 949 CCGCGCCGACTGCGACAGCTGCTAGCGGGCATGATCGAGGTGAGACGCGCGCCCT 1008
Qy 597 CACGTAGAGCACATTCAGAAAAACATATGATGATATAGGGGGCAAGAAATGAGCCGGAACC 656
Db 1009 CACGTAGAGCACATTCAGAAAAACATATGATGATATAGGGGGCAAGAAATGAGCCGGAACC 1068
Qy 657 AGAGCAGCCCATTTCTCGAAGGTGCAGATCCGCTCGCTGCCAGCTGTGAGGACATCGA 716
Db 1069 AGAGCAGCCCATTTCTCGAAGGTGCAGATCCGCTCGCTGCCAGCTGTGAGGACATCGA 1128
Qy 717 CCGCGAGCTGCTGACAGCATGCACTCAGTGGGCTGCTTCCGAGACCGCAACAGCTGCT 776
Db 1129 CCGCGAGCTGCTGACAGCATGCACTCAGTGGGCTGCTTCCGAGACCGCAACAGCTGCT 1188
Qy 777 GCAGGACCTGCTGCTCGAGGAGAGAACAGAGAGATGATTTACTTCTCTCTCGGA 836
Db 1189 GCAGGACCTGCTGCTCGAGGAGAGAACAGAGAGATGATTTACTTCTCTCTCGGA 1248
Qy 837 CCGGAAAGAAAGTACCCGAGCAGAGGATGAGGACCTGCCCCCGGAAACGAGATAGA 896
Db 1249 CCGGAAAGAAAGTACCCGAGCAGAGGATGAGGACCTGCCCCCGGAAACGAGATAGA 1308
Qy 897 CCTTCCCGGAGCGTGTGACTCCCGATGCTAAACCGGACGCGCAAGCGGGCCAGA 956
Db 1309 CCTTCCCGGAGCGTGTGACTCCCGATGCTAAACCGGACGCGCAAGCGGGCCAGA 1368
Qy 957 ACGCAATTCATGAGGTGCTCAGCGTGA CCGACGCGGCTCCCGGTGCTGCGCGGCG 1016
Db 1369 ACGCAATTCATGAGGTGCTCAGCGTGA CCGACGCGGCTCCCGGTGCTGCGCGGCG 1428
Qy 1017 GGCATTTGAGATGGCCACGACCGCCAGAGGTCTCGGTCCATCAGCGGTGCTCTCAGG 1076
Db 1429 GGCATTTGAGATGGCCACGACCGCCAGAGGTCTCGGTCCATCAGCGGTGCTCTCAGG 1488
Qy 1077 CTTTTCACAGCCCACTCAGCAGCCCCCGGGTGA CCCCCTCACAAAGGGGCGAG 1136
Db 1489 CTTTTCACAGCCCACTCAGCAGCCCCCGGGTGA CCCCCTCACAAAGGGGCGAG 1548
Qy 1137 TCCCTTCCCCCA CCGCCAAAGGGGACACTGTGTCCA CAGCCAAAGGAGAGCCCGGCTGGCAC 1196
Db 1549 TCCCTTCCCCCA CCGCCAAAGGGGACACTGTGTCCA CAGCCAAAGGAGAGCCCGGCTGGCAC 1608
Qy 1197 GCTCAACCCCA CCGCCCGCTG CAGCCCCAGCGTTCGAGGGGTGCTTGGAGGGGCGGGCT 1256
Db 1609 GCTCAACCCCA CCGCCCGCTG CAGCCCCAGCGTTCGAGGGGTGCTTGGAGGGGCGGGCT 1668
Qy 1257 CAATCTCATCAAGAAACAGCTTTCTGGGCTCA CCGCGCTCCACCGCGGAAATGCAAGT 1316
Db 1669 CAATCTCATCAAGAAACAGCTTTCTGGGCTCA CCGCGCTCCACCGCGGAAATGCAAGT 1728
Qy 1317 TCCGAGCGCGGAGAGATGTCCAACCTGACA CAGAGTCTGCTCCCGAGAGTGGCGGAAGAA 1376
Db 1729 TCCGAGCGCGGAGAGATGTCCAACCTGACA CAGAGTCTGCTCCCGAGAGTGGCGGAAGAA 1788
Qy 1377 GTCTGTGTTTGGAACTTTCATCAGCTTGGAGAGGAGGACAGATCTTCTGTGTCTCAA 1436
Db 1789 GTCTGTGTTTGGAACTTTCATCAGCTTGGAGAGGAGGAGGACAGATCTTCTGTGTCTCAA 1848
Qy 1437 AGCAAAACCTCTGAGCTTCCATCAAGGCTGACATCTGTCAGCGCTTCTCTGTCTGATTCACAG 1496
Db 1849 AGCAAAACCTCTGAGCTTCCATCAAGGCTGACATCTGTCAGCGCTTCTCTGTCTGATTCACAG 1908
Qy 1497 TCTAGCCACAGCGTCTCATCTCCCAAAACAGAGCTTTCGGGGCGAGGTACAAAGCCACGGGGGG 1556


```
Db 1493 TCCATCAGAACAGCTTTCTGGGCTACCCCGCTTCCACCGCGGAACCTCAAGTTCCG 1552
Qy 1321 AGCGCGAGGAGATGTTCAACCTGACACGAGATGCTGCCAGAGCTGGGGAAGATCC 1380
Db 1553 ACGCCGAGGAGATGTTCAACCTGACACGAGATGCTGCCAGAGCTGGGGAAGATCC 1612
Qy 1381 TGGTTTGGGAACCTTCAATCAGCTCGAGAGGAGAGAGATCTTCTGTTGTTCAATCAAGAC 1440
Db 1613 TGGTTTGGGAACCTTCAATCAGCTCGAGAGGAGAGAGATCTTCTGTTGTTCAATCAAGAC 1672
Qy 1441 AAACCTCTGAGCTCCATCAAGGCTGACATCGTGACAGCTTCTCTGTCGATTCCTCAGTCTC 1500
Db 1673 AAACCTCTGAGCTCCATCAAGGCTGACATCGTGACAGCTTCTCTGTCGATTCCTCAGTCTC 1732
Qy 1501 AGCCACAGCTCATCTCCAAACGAGCTTCCGGCGGAGTACAGGCAACGGGGGGGCA 1560
Db 1733 AGCCACAGCTCATCTCCAAACGAGCTTCCGGCGGAGTACAGGCAACGGGGGGGCA 1792
Qy 1561 GCCGTGTTCCAGAACCGGTCAGGTTCCAGGTTGATATCACCTACACGAGGGGTGGGGAG 1620
Db 1793 GCCGTGTTCCAGAACCGGTCAGGTTCCAGGTTGATATCACCTACACGAGGGGTGGGGAG 1852
Qy 1621 GCGCAGAGGAGAACGGCATCTACTTCCTGTCACCTTACCTGCTCAGGCCCCCAGCGGT 1680
Db 1853 GCGCAGAGGAGAACGGCATCTACTTCCTGTCACCTTACCTGCTCAGGCCCCCAGCGGT 1912
Qy 1681 CGCTTCAGAGGGGTGGTGGAGACCATCCAGGCCAGCTGCTGAGCACACAGACCCGCT 1740
Db 1913 CGCTTCAGAGGGGTGGTGGAGACCATCCAGGCCAGCTGCTGAGCACACAGACCCGCT 1972
Qy 1741 GCGGCCAGCACTTGTGACAGACCAAC 1766
Db 1973 GCGGCCAGCACTTGTGAGAACCCCC 1998

RESULT 11
US-10-283-247-4
; Sequence 4, Application US/10283247
; Publication No. US20030119037A1
; GENERAL INFORMATION:
; APPLICANT: NEELAM, Beena et al.
; TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC
; ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES
; TITLE OF INVENTION: THEREOF
; FILE REFERENCE: CL001304
; CURRENT APPLICATION NUMBER: US/10/283,247
; CURRENT FILING DATE: 2002-10-30
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 1911
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-283-247-4

Query Match 92.2%; Score 1685.2; DB 14; Length 1911;
Best Local Similarity 96.2%; Pred. No. 0;
Matches 1763; Conservative 0; Mismatches 3; Indels 56; Gaps 1;

Qy 1 ATGAAGGTGGAGCGGAGATCCGATCTCTGAAGCTCATTTGAGCACCCCAAGTCTCTAAAG 60
Db 1 ATGAAGGTGGAGCGGAGATCCGATCTCTGAAGCTCATTTGAGCACCCCAAGTCTCTAAAG 60
Qy 61 CTGACGAGCTTTATGAACCAAAATATTTGTACCTGTTGCTAGACACGTTGTCAGGT 120
Db 61 CTGACGAGCTTTATGAACCAAAATATTTGTACCTGTTGCTAGACACGTTGTCAGGT 120
Qy 121 GGTGAGCTTCTGATCTACCTGTTGAAGAGGGAGGCTGACCCCTAAGGAGGCTCGGAAG 180
Db 121 GGTGAGCTTCTGATCTACCTGTTGAAGAGGGAGGCTGACCCCTAAGGAGGCTCGGAAG 180
Qy 181 TTCTTCCCGCAGATCATCTCTCGCTGGAATTTCTGCCACAGCACTCCATATGCCACAGG 240
Db 181 TTCTTCCCGCAGATCATCTCTCGCTGGAATTTCTGCCACAGCACTCCATATGCCACAGG 240
```

```
Db 181 TTCTTCCCGCAGATCATCTCTCGCTGGAATTTCTGCCACAGCACTCCATATGCCACAGG 240
Qy 241 GATCTGAAAACCTGAAAACCTCTCTGCTGGAGAGAGAAACAACATCCGCACTGGAGACTTT 300
Db 241 GATCTGAAAACCTGAAAACCTCTCTGCTGGAGAGAGAAACAACATCCGCACTGGAGACTTT 300
Qy 301 GGCATGCGCTCCCTGCAAGCTTGGCGACAGCTGTTGGAGACCAAGCTGTGGTCCCCCCAC 360
Db 301 GGCATGCGCTCCCTGCAAGCTTGGCGACAGCTGTTGGAGACCAAGCTGTGGTCCCCCCAC 360
Qy 361 TACGCTGCCCCGAGATGATCCGGGGGAGAAATATGACGCGCGGAGGCGAGCTGTGG 420
Db 361 TACGCTGCCCCGAGATGATCCGGGGGAGAAATATGACGCGCGGAGGCGAGCTGTGG 420
Qy 421 AGCTGGGGGCTCATCTGTTTCCGCTTGGCTGGGGGCTCTGCCCTTCGACGATGACAAC 480
Db 421 AGCTGGGGGCTCATCTGTTTCCGCTTGGCTGGGGGCTCTGCCCTTCGACGATGACAAC 480
Qy 481 TTCCGACAGCTGCTGGAGAAAGTGAAGCGGGGCGTGTTCACATGCGCGCATTTATCCCG 540
Db 481 TTCCGACAGCTGCTGGAGAAAGTGAAGCGGGGCGTGTTCACATGCGCGCATTTATCCCG 540
Qy 541 CCCGACTGCGCAGAGTCTGCTTACGCGGCGATGATCGAGGTGAGACCGCGCCCTCACG 600
Db 541 CCCGACTGCGCAGAGTCTGCTTACGCGGCGATGATCGAGGTGAGACCGCGCCCTCACG 600
Qy 601 CTAGAGCACATTCAGAAAACAACATATGATATATAGGGGGCAAGAAATGAGCCCGAACAGAG 660
Db 601 CTAGAGCACATTCAGAAAACAACATATGATATATAGGGGGCAAGAAATGAGCCCGAACAGAG 660
Qy 661 CAGCCCAATTCCTCGCAAGGTGAGATCCGCTCGCTGCCAGCCTGGAGGACATCGACCCC 720
Db 661 CAGCCCAATTCCTCGCAAGGTGAGATCCGCTCGCTGCCAGCCTGGAGGACATCGACCCC 720
Qy 721 GAGCTGCTGACAGATGCACTCACTGGGGTGTCTCCGAGACCGCAACAAGCTGTGTCAG 780
Db 721 GAGCTGCTGACAGATGCACTCACTGGGGTGTCTCCGAGACCGCAACAAGCTGTGTCAG 780
Qy 781 GACCTGCTGTCGAGGAGGAGAACACAGGAGAAATGATTTACTTCTCTCTCTGGACCGG 840
Db 781 GACCTGCTGTCGAGGAGGAGAACACAGGAGAAATGATTTACTTCTCTCTCTGGACCGG 840
Qy 841 AAAGAAAGGTACCCGACGACGAGGAGATGAGACCTGCCCCCGGCAACGAGATAGACCT 900
Db 841 AAAGAAAGGTACCCGACGACGAGGAGATGAGACCTGCCCCCGGCAACGAGATAGACCT 900
Qy 901 CCGCGAAGCGGTGTGAGTCTCCCGATGCTGAAACGCGGACCGGCAAGCGCGCGGACAGCG 960
Db 901 CCGCGAAGCGGTGTGAGTCTCCCGATGCTGAAACGCGGACCGGCAAGCGCGCGGACAGCG 960
Qy 961 AAATCCATGAGGTGCTCAGCGTGAACGCGGCGGCTTCCCGGCTGCTGCGCGCGGCGCC 1020
Db 961 AAATCCATGAGGTGCTCAGCGTGAACGCGGCGGCTTCCCGGCTGCTGCGCGCGGCGCC 1020
Qy 1021 ATTGAGATGCGCCAGCACCGCCAG- 1044
Db 1021 ATTGAGATGCGCCAGCACCGCCAGAGTAAAGCAATGTTCAAGTAAAGCCTTGGATATCGCT 1080
Qy 1045 -----AGTCTCGTTCATCAGCGGTGCTCTCTCA 1074
Db 1081 GAGGCCCATCCCCAATTCAGCAAGAAAGACAGGCTCTCGGTCCATCAGCGGTGCTCTCTCA 1140
Qy 1075 GGCCTTTCCACACGACCCCACTCAGCAGCCCCGGGTGACCCCTCAACCCCTCAACCAAGGGGC 1134
Db 1141 GGCCTTTCCACACGACCCCACTCAGCAGCCCCGGGTGACCCCTCAACCCCTCAACCAAGGGGC 1200
Qy 1135 AGTCCCTTCCCAACCCCAAGGGGACACCTGTCCACACGCGCAAGAGAGAGCCCGGTGGC 1194
Db 1201 AGTCCCTTCCCAACCCCAAGGGGACACCTGTCCACACGCGCAAGAGAGAGCCCGGTGGC 1260
Qy 1195 ACGCCCAACCCCAAGCCCGCTCCAGCCCCAGCGTCCGAGGGGTGCCCTGGAGGGCGCGG 1254
Db 1261 ACGCCCAACCCCAAGCCCGCTCCAGCCCCAGCGTCCGAGGGGTGCCCTGGAGGGCGCGG 1320
```


QY 1255 CTCAACTCCATCAAGAACAGCTTTCTGGGCTCAACCCGCTTCCACCGCGGAAACTGCAA 1314
Db
1321 CTCAACTCCATCAAGAACAGCTTTCTGGGCTCAACCCGCTTCCACCGCGGAAACTGCAA 1380
QY 1315 GTTCCGACGCGGAGAGATGCTCAACCTGACACAGAGTGCTCCCAAGAGCTGGCGAAG 1374
Db 1381 GTTCCGACGCGGAGAGATGCTCAACCTGACACAGAGTGCTCCCAAGAGCTGGCGAAG 1440
QY 1375 AAGTCTCTGTTGGGAACTTCAATCAGCTTGGAGAGGAGGAGAGATCTTGGTGTATC 1434
Db 1441 AAGTCTCTGTTGGGAACTTCAATCAGCTTGGAGAGGAGGAGAGATCTTGGTGTATC 1500
QY 1435 AAAGACAAACCTCTGAGCTTCAATCAAGCTGACATCGTGACGCTTCTCTGCGAATTCCT 1494
Db 1501 AAAGACAAACCTCTGAGCTTCAATCAAGCTGACATCGTGACGCTTCTCTGCGAATTCCT 1560
QY 1495 AGTCTCAGCACAGCGTCTATCTCCAAACAGAGCTTCCGCGGCGAGTACAAGGCCACGCGG 1554
Db 1561 AGTCTCAGCACAGCGTCTATCTCCAAACAGAGCTTCCGCGGCGAGTACAAGGCCACGCGG 1620
QY 1555 GGGGAGCGCGTCTCAGAGCGGCTCAAGCTGACATCGTGACGCTTCTCTGCGAATTCCT 1614
Db 1621 GGGGAGCGCGTCTCAGAGCGGCTCAAGCTGACATCGTGACGCTTCTCTGCGAATTCCT 1680
QY 1615 GGGGAGCGCGTCTCAGAGCGGCTCAAGCTGACATCGTGACGCTTCTCTGCGAATTCCT 1674
Db 1681 GGGGAGCGCGTCTCAGAGCGGCTCAAGCTGACATCGTGACGCTTCTCTGCGAATTCCT 1740
QY 1675 AGCGTCTGCTTCAAGAGGCTGTGAGACCATCCAGGCGGAGTCAAGGCCACACGAC 1734
Db 1741 AGCGTCTGCTTCAAGAGGCTGTGAGACCATCCAGGCGGAGTCAAGGCCACACGAC 1800
QY 1735 CGCGCTGCGGCGGAGCTTTCAGACACAC 1766
Db 1801 CGCGCTGCGGCGGAGCTTTCAGACACAC 1832

RESULT 12
US-10-116-326-5
; Sequence 5, Application US/10116326
; Publication No. US20030166889A1
; GENERAL INFORMATION:
; APPLICANT: Turner, C. Alexander Jr.
; APPLICANT: Mathur, Brian
; APPLICANT: Fridele, John
; TITLE OF INVENTION: No. US20030166889A1 Human Kinases and Polynucleotides Encoding
; FILE REFERENCE: LEX-0332-USA
; CURRENT APPLICATION NUMBER: US/10/116,326
; CURRENT FILING DATE: 2002-04-04
; PRIOR APPLICATION NUMBER: US 60/282,036
; PRIOR FILING DATE: 2001-04-06
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 5
; LENGTH: 2112
; TYPE: DNA
; ORGANISM: homo sapiens
US-10-116-326-5

Query Match 46.6%; Score 851.4; DB 12; Length 2112;
Best Local Similarity 68.9%; Pred. No. 5e-224;
Matches 1356; Conservative 0; Mismatches 416; Indels 195; Gaps 6;

QY 1 ATGAGGTGAGCGGAGATCGGATCTCGAGCTCAATGAGCACCCCACTCTCTAAG 60
Db 1 ATGAGGTGAGCGGAGATCGGATCTCGAGCTCAATGAGCACCCCACTCTCTAAG 60
QY 61 CTGACGAGCTTATGAAACAAATAATTTGTACTGCTGTAGAACACCTGTGAGT 120
Db 61 CTGACGAGCTTATGAAACAAATAATTTGTACTGCTGTAGAACACCTGTGAGT 120
QY 121 GGTGAGCTTCTGACTACCTGCTGAGAGAGGAGGCTGACGCTTAAGAGGCTCGGAAG 180

Db 121 GGTGAGCTTCTGACTACCTGCTGTAAGAGAGGAGAGCTGAGCGCCCAAGGCGCGGAAG 180
QY 181 TTCTTCCGAGATCATCTCTGGCTGAGATTTCTGGCCACAGCCTCATATGACAGG 240
Db 181 TTCTTCCGAGATCATCTCTGGCTGAGATTTCTGGCCACAGCCTCATATGACAGG 240
QY 241 GATCTGAAACCTGAAACCTCTGCTGAGAGAGAACCAACATCCGATTCGAGACTTT 300
Db 241 GATCTGAAACCTGAAACCTCTGCTTGGATGAGAAACCAACATCCGATTCGAGACTTT 300
QY 301 GGATGCGCTCTCTGAGGTTGGCAGCAGCTTGTGGAGACCAAGCTGTGGGTCCCCC 360
Db 301 GGATGCGCTCTCTGAGGTTGGCAGCAGCTTGTGGAGACCAAGCTGTGGGTCCCCC 360
QY 361 TACGCTCTCCCGAGGTTGATCCGCGGAGAGATGATACCGCGGAGAGGCGAGTGTGG 420
Db 361 TACGCTCTCTGAGGTTGATTAAGGGGGAATAATATGATGCGCGCGGAGAGATGTGG 420
QY 421 AGCTGCGGCTCATCTCTGCTTGGCTTCTGCTGGGGGCTCTGCGCTTCGACGATGACAAC 480
Db 421 AGCTGCGGCTCATCTCTGCTTGGCTTCTGCTGGGGGCTCTGCGCTTCGACGATGACAAC 480
QY 481 TTGCGACAGCTGCTGAGAGAGGTTGAAGCGGGGCTGTTCACATGCGCGACTTTATCCCG 540
Db 481 TTGCGACAGCTGCTGAGAGAGGTTGAAGCGGGGCTGTTCACATGCGCGACTTTATCCCG 540
QY 541 CCGACCTGCGGAGCTGCTACCGGGGATGATGAGGTGAGCGCGGACCGCGCTCAGG 600
Db 541 CCGACCTGCGGAGCTGCTACCGGGGATGATGAGGTGAGCGCGGACCGCGCTCAGG 600
QY 601 CTAGAGCACATTCAGAAACACATATGATGAGGGGCAATGATGAGGGGCGGAAAGGCTCAGT 656
Db 601 CTAGAGCACATTCAGAAACACATATGATGAGGGGCAATGATGAGGGGCGGAAAGGCTCAGT 656
QY 657 --AGAGCAGCGCTTCTGCGAAGGTGAGATCGCTGCTGCGCGAGCTTGGAGGACATC 714
Db 657 --AGAGCAGCGCTTCTGCGAAGGTGAGATCGCTGCTGCGCGAGCTTGGAGGACATC 714
QY 715 GACCCGAGCTGCTGAGACAGCATCTACTGGGCTGCTCCGAGACCGGCAACAGCTG 774
Db 715 GACCCGAGCTGCTGAGACAGCATCTACTGGGCTGCTCCGAGACCGGCAACAGCTG 774
QY 775 CTGAGGACCTGCTGCTCGAGGAGAGAACAGAGAGAGATGATTTACTCTCTCTCTG 834
Db 775 CTGAGGACCTGCTGCTCGAGGAGAGAACAGAGAGAGATGATTTACTCTCTCTCTG 834
QY 835 GACCGAAAGAAAGTACCCGAGCAGGAGATGAGGACCTGCGCGCGCGGAAACGAGATA 894
Db 835 GACCGAAAGAAAGTACCCGAGCAGGAGATGAGGACCTGCGCGCGCGGAAACGAGATA 894
QY 895 GACCTCTCCCGAAGCTGTGGACTCCCGGATGCTGAACCGGCGAGCGGAGCGGCGGCA 954
Db 895 GACCTCTCCCGAAGCTGTGGACTCCCGGATGCTGAACCGGCGAGCGGAGCGGCGGCA 954
QY 901 GACCCCGCGGAGCGGTGATTTCTCCATGCTGAGCGCTGACGCGGAGCGGCGACCA 960
Db 901 GACCCCGCGGAGCGGTGATTTCTCCATGCTGAGCGCTGACGCGGAGCGGCGACCA 960
QY 955 GAAAGCAATCCATGAGGCTGCTCAGCGTGAAGCA-----CGGCGGCTCCCGGCTG 1005
Db 955 GAAAGCAATCCATGAGGCTGCTCAGCGTGAAGCA-----CGGCGGCTCCCGGCTG 1005
QY 1006 CCGCGGCGGCGGCTTGAAGATGCGGCGAGCAAGCGGCGAGGCTTCTGGTTCATCAGCGGT 1065
Db 1006 CCGCGGCGGCGGCTTGAAGATGCGGCGAGCAAGCGGCGAGGCTTCTGGTTCATCAGCGGT 1065
QY 1021 CCCACCGGCGGCTTGGAGATGCGGCGAGCAGCAGGAGATCCCGTAGCGTCAAGTGA 1080
Db 1021 CCCACCGGCGGCTTGGAGATGCGGCGAGCAGCAGGAGATCCCGTAGCGTCAAGTGA 1080
QY 1066 GCTCTCTCAGGCTTTCACCGGCGGCTCAGGAGCGGCGGCGG-----1108
Db 1066 GCTCTCTCAGGCTTTCACCGGCGGCTCAGGAGCGGCGGCGG-----1108
QY 1081 GCCTCCAGGCTCTGCTCTCCAGCGCTTAAGCAGCGGCGGCGGCTTCTTCTCTTT 1140
Db 1081 GCCTCCAGGCTCTGCTCTCCAGCGCTTAAGCAGCGGCGGCGGCTTCTTCTCTTT 1140
QY 1109 -----1108
Db 1109 -----1108
QY 1141 TCACCGGAGCGGCGGCTGAGATGAGGCTCGAGCGGCGGCGGCTCCCGACTTCCAAAACG 1200
Db 1141 TCACCGGAGCGGCGGCTGAGATGAGGCTCGAGCGGCGGCGGCTCCCGACTTCCAAAACG 1200
QY 1109 -----TGACCCCTCAACCC 1122
Db 1109 -----TGACCCCTCAACCC 1122

Db 1201 CAGAGCGTCTCTCGGGGCCCCAGAGGTGGGGCGCGGGGAGCAGCCCCCGCCCCC 1260
Qy 1123 TCACCAAGGGGAGTCCCTCC-----ACC 1149
Db 1261 AGTCCCGCTCCACACCCCTGCGGGGCCCCCAGGCTCCCGCGCTCTCTGCGGGACC 1320
Qy 1150 CCCAAGGGGACACTGTCCACACGCCAAAGAGAGCCGGCTGCGACGCCCAACCCACAG 1209
Db 1321 CCCTTGACTCGCTCTGCAACGCCCGGGGCGAGTCCACCGGACCCCGGGGACAACA 1380
Qy 1210 CCCCCGTCAGCCCC-----AGCGTGAGGGGTGCCCTGAGGGGCGGCTCAACTCC 1263
Db 1381 CCACCCCCAGCCCCGGGTGGGTGCGGGGAGCGCGCTGAGGAGTCTCAACTCC 1440
Qy 1264 ATCAAGACAGCTTCTGGGTCTACCCCGCTTCCACCGCGGAACTGCAAGTTCGACG 1323
Db 1441 ATCCGCAACAGCTTCTGGGTCCCTCGCTTTCACCGGCGAAGTGCAGGTCCCTACC 1500
Qy 1324 CCGGAGGAGATGTCCAACTGACACCAAGTCTGTCGCCCGGAGCTGGGCAAGTCTCTGG 1383
Db 1501 GCTGAGGAGATGTCCAGTTCAGCGCAGAGTCTCCCGGAGCTGGCAACGCTCTCTGG 1560
Qy 1384 TTGGGAATTCATCAGCCTGGAGAGGAGAGATCTTCTGGTGTATCACTACAAGACAAA 1443
Db 1561 TTGGGAATTCATCTCTGACAAAGAGACAAATATTTCTCTGTCTAAAGACAAA 1620
Qy 1444 CCTCTGAGCTCCATCAAGCTGACATCGTGCAGCGCTTCTGTGATTCACAGTCTCAGC 1503
Db 1621 CCTCTCAGCAGCATCAAGCAGACATCTCCATGCTTCTGTGATCCCCAGGCTGAGT 1680
Qy 1504 CACAGCGTCATCTCCAAACAGCTTCCGGGCGGAGTACAAAGGCCACCGGGGGGCGAGCC 1563
Db 1681 CACAGTGTCTCAGACACAGCTTTCAGGCGAGTACAGGCGAGTGGCGGCCCTCC 1740
Qy 1564 GTGTTCCAGAACCGGTCAAGTTCAGGTTGATATCACTACACGAGGTT----- 1614
Db 1741 GTCTTCCAAAGCCGTCGCTTCCAGTGGACATCAGCTCTCTGAGGGTCCAGAGGCC 1800
Qy 1615 -----GGGGAGCGCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1665
Db 1801 TCCCGCGAGCGGAGCGCAGCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1860
Qy 1666 TCAGGCGCCAGCGCTCGCTTCAAGAGGGTGGTGGAGACCATCCAGGCCAGCTGTGAGC 1725
Db 1861 TCGGGTCCAGCGCTCGGTTCAGCGAGTGGTGGAGACCATCCAGSCACAGCTCTGAGC 1920
Qy 1726 ACACGAGCCCGCTGCGGCCAGCAGCTTGTGAGACACCACTAACTG 1772
Db 1921 ACTCATGACAGCCCTCCGTGCGAGCCCTGGCAGACGAGAGAGAGAGAGAGAGAGAG 1967

RESULT 13

US-10-116-326-1

; Sequence 1, Application US/10116326

; Publication No. US2003016689A1

; GENERAL INFORMATION:

; APPLICANT: Mathur, C. Alexander Jr.

; APPLICANT: Friddle, Brian

; TITLE OF INVENTION: No. US2003016689A1 Human Kinases and Polynucleotides Encoding

; FILE REFERENCE: LEX-0332-USA

; CURRENT APPLICATION NUMBER: US/10/116,326

; CURRENT FILING DATE: 2002-04-04

; PRIOR APPLICATION NUMBER: US 60/282,036

; PRIOR FILING DATE: 2001-04-06

; NUMBER OF SEQ ID NOS: 6

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 1

; LENGTH: 2337

; TYPE: DNA

; ORGANISM: homo sapiens

US-10-116-326-1

Query Match 46.6%; Score 851.4; DB 12; Length 2337;
Best Local Similarity 68.9%; Pred. No. 5.2e-224;
Matches 1356; Conservative 0; Mismatches 416; Indels 195; Gaps 6;
Qy 1 ATGAAGGTGGAGCGGAGATCGGATCTCTGAAGCTCATTTAGACACCCACGTCCTTAAAG 60
Db 226 ATGAAGGTGGAGCGGAGATCGGATCTCTGAAGCTCATCGAAACACCCACATGTCTCTCAAG 285
Qy 61 CTGCAGAGCTTTATGAAACAAATAATTTGTACTGTTGCTAGAACACGTTGTCAGGT 120
Db 286 CTCACGAGCTCTACGAGAACAAATAATTTGTACTGTTGCTGAGACACGTTCTCGGG 345
Qy 121 GGTGAGCTCTTTCGACTTACCTGTGTGAAGAGGGGAGGCTGACGCTTAAGGAGGCTCGGAAG 180
Db 346 GGTGAGCTATTTCGACTTACCTGTGTGAAGAGGGGAGAGTGCAGCCCAAGGAGGCCGGAAG 405
Qy 181 TTCTTCCGGGAGATCATCTCTGGCTTGGACTTCTGCGACAGCCACTTCCATATGCCACAGG 240
Db 406 TTCTTCCGGGAGATTTGTCTGGCTTGGACTTCTGCGACAGCTTCTCCATCTGCCACAGA 465
Qy 241 GATCTGAAACCTGAAACCTCTCTGCTGGAGAGAGAAACAACATCCGCACTCGCAGACTTT 300
Db 466 GACCTAAAGCCCGAGAACCTCTCTTTGGATGAGAAAAACAACATCCGCACTTGCAGACTTC 525
Qy 301 GGCATGGCGTCCCTGCAAGTGTGGCAGCAGCTCTTGGAGACCAAGCTGTGGGTCCCCCAC 360
Db 526 GGCATGGCGTCCCTGCAAGTGTGGGAGACGCTCTCTGGAGACCAAGCTGTGGGTCCCCCAT 585
Qy 361 TAGCCTGCCCGGAGTGTATCCGGGGGAGAGATATGACGGCCGGAAGGGCGGAGTGTGG 420
Db 586 TATGCGTGTCCAGAGGTGATTAAGGGGGAAAAATATGATGGCCCGCGGGCAGACATGTGG 645
Qy 421 AGCTTGGCGGTCTATCTCTGCTGCTGGGCTCTGCTGGGCTCTGCTTTCGACGATGACAAC 480
Db 646 AGCTTGGAGTCTATCTCTGCTGCTGGGCTCTGCTGGGCTCTGCTTTCGACGACAC 705
Qy 481 TTGCGACAGCTGTGGAGAGAGTGAAGGGGGGTGTTCACATGCGCGCATTTATCCCG 540
Db 706 CTCGCCAGCTGTGGAGAGGTGAACGGGGGTCTTCCACATGCCCCACTTCATTCTCT 765
Qy 541 CCGAGTCCAGAGTCTGTACGGGGCATGATCGAGGTGAGACGGCGCGCGCTCAGC 600
Db 766 CCAGATTGCCAGAGCTCTCTGAGGGGAAATGATCGAAGTGGAGCCCGGAAAAAGGCTCAGT 825
Qy 601 CTAGAGCACATTCAGAAACACATATGTTATATAGGGGGCAAGATGAGCCCGAACCC--- 656
Db 826 CTGAGAGCAAAATTCAGAAACATCTCTGTGTACTAGCGGGGAAACACAGAGCCAGACCCGTGC 885
Qy 657 --AGAGCAGCCCATTTCTCGAAAGGTGACAGATCCGCTGCTGCCCGCAGCTTGGAGGACATC 714
Db 886 CTGGAGCCAGCCCTGCGCCCGGGTAGCCATGCGGAGCTTGCATCCAAACGGAGAGCTG 945
Qy 715 GACCCGAGCTGTGGAGACATGCACTCACTGGGCTGCTTCGAGAGCCGCAACAGCTG 774
Db 946 GACCCGAGCTCTTAGAGAGCATGGCATCACTGGGCTGCTTCAGGGGACCCCGAGAGCTG 1005
Qy 775 CTGCAAGACCTGTGTCGAGGAGAGAGAAACAGGAGAGATGATTTACTTCTCTCTCTG 834
Db 1006 CATCGGAGCTGCGCAGTGTAGGAGAGAGAACAGAGAAAGATGATATATATCTGCTTTG 1065
Qy 835 GACCGGAAAGAGGTATCCGAGCAGGAGGATGAGACCTGCCCCCGGAAACGAGATA 894
Db 1066 GATCGGAAGGAGCGGTATCCAGCTGTGAGGACAGGACCTGCTCCCGGAATGATGTT 1125
Qy 895 GACCTCCCGGAGAGCTGTGAGTCCCGATGCTGAACCGGCAACCGCAGACGGCGGCCA 954
Db 1126 GACCCCGGAGAGCGTGTGATTTCTCCATGCTGAGCCGTACCGGGAACGGCGGACCA 1185
Qy 955 GAACGCAAAATCCATGAGGAGTGTCTCAGCGTGAACGGA-----CGCGGGTCCCCCGGTG 1005
Db 1186 GAGCGGAAGTCCATGGAAGTCTCTGAGCATCACCGATCCCGGGGGTGTGGCTCCCTGTGA 1245
Qy 1006 CCTGCGCGGGGCCCATTTGAGATGGCCCAAGCAAGCGCCAGAGGTCTCGGTTCATCAGCGGT 1065

```

1246 CCCACCCGAGCGGCTTGGAGATGCCACGACGACGAGATCCGTAGGTCAAGTGA 1305
1066 GCGCTCTCAGGCTTTCACGAGCCACTCAGCAGCCCGCGG----- 1108
1306 GCGCTCAGCGGCTGTCTCTCCAGCCCTCTAAGCAGCCCAAGGAGTCCGGTCTTTTCTTT 1365
1109 ----- 1108
1366 TCACCGGAGCCGGGGCTGGAGATGAGGCTCAGCGGGGGGCTCCCGACATTCGAAAGG 1425
1109 -----TGACCCCTCACCCC 1122
1426 CAGACGCTGCTTCTCGGGGCCCCAGGGGTGGGGGCGCGGGGAGCAGCCCCCGCCCCC 1485
1123 TCACCAAGGGGAGTCCCTCTCC-----ACC 1149
1486 AGTGCCCGTCCACACCCCTGCCCCGCCCCCAGGCTCCCGCGTCTCTTGGCGGGACC 1545
1150 CCCAAGGGGACACCTGTCTCACACGCGCAAGGAGAGCCCGGCTGGCAGCCCAACCCACG 1209
1546 CCGTTGCACTCGCTCTGACAGCGCCCGGGCCAGTCCACCGGACCCCGGGGACAACA 1605
1210 CCCCCGCTCAGCCCC-----AGCGTCGAGGGGTGCCCTGGAGGGCGGGCTCAACTCC 1263
1606 CCACCCCCCAGCCCGGGGTGGCGTCTGGGGGAGCGCTGGAGGAGTCTCTCAACTCC 1665
1264 ATCAGAGACAGCTTCTGGGCTCACCCGCTTCCACCGCGGCAAACTCGAGTCCGAGG 1323
1666 ATCCGCAACAGCTTCTGGGCTCCCTCTGCTTTTCCACCGGCGCAAGTGCAGGTCCCTACC 1725
1324 CCGGAGGAGATGCCAACTGACACAGAGTCTGCCAGAGCTGGCGAAAGTCTCTGG 1383
1726 GCTGAGGAGATGTCAGCTTGAGCCAGAGTCTTCCCGGAGCTGGCAAAAGCTCTCTGG 1785
1384 TTGGAAGTTCATCAGCTTGAGAGGAGGAGAGTCTTGGTGGTTCATCAAGACAAA 1443
1786 TTGGGAAGTTCATCTCTTGGACAAAGAAACAATATTCCTGCTGCTAAAGGACAAA 1845
1444 CTTCTGAGTCCATCAAGGCTCAGATCTGTCAGCGCTTCTGTCGATTCCTGCTCAGC 1503
1846 CTTCTCAGAGATCAAGAGACATCTGTCATGCTTCTGTCATCCCGAGCTGAGT 1905
1504 CACAGCTCATCTCCCAAACGAGCTTCCGGGCGGAGTACAGGCCACGCGGGGCGCAGCC 1563
1906 CACAGTGTCTCTACAGACAGCTTCCAGGCGGAGTACAGGCCAGTGGCGGCCCTCC 1965
1584 GTGTTCCAGAGCCGGTCAAGTTCAGGTTGATATCACTTACAGCGAGT----- 1614
1966 GTCTTCCAAAGCCGCTCCGCTTCCAGGTGGACATCAGCTCTCTGAGGGTCCAGAGGCC 2025
1615 -----GGGGGCGGCGAGAGGAGAACCGCATCTACTCGTCACCTTCCACCTGCTC 1665
2026 TCCCGCGAGCGGAGCGGAGCGGAGGTGGTGGCATCTACTCGTCACCTTCACTCTCATC 2085
1666 TCAGGCCCGAGCGCTCGCTTCAAGAGGGTGGTGGAGACCATCCAGGCCAGCTGCTGAGC 1725
2086 TCGGGTCCAGCGCTCGGTTCAAGCGAGTGGTGGAGACCATCCAGGCACAGCTCTCTGAGC 2145
1726 ACACGAGCCCGCTCGGCGCCAGCAGCTTGTGAGACACCATCTG 1772
2146 ACTCATGACGAGCCCTCGGTCAGGCGCTTGGCAGAGGAGAACGG 2192

```

```

RESULT 14
US-10-354-358-91
; Sequence 91, Application US/10354358
; Publication NO. US20030157082A1
; GENERAL INFORMATION:
; APPLICANT: Millennium Pharmaceuticals, Inc
; APPLICANT: Hunter, John Joseph
; APPLICANT: Macbeth, Kyle J.
; APPLICANT: Tsai, Fong-Ying

```

```

; APPLICANT: Lesoon, Andrea
; APPLICANT: Lightcap, Eric S.
; APPLICANT: Williamson, Mark
; APPLICANT: Rudolph-Owen, Laura A.
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR TREATING
; TITLE OF INVENTION: CANCER USING 140, 1470, 1686, 2089, 2427, 3702, 5891, 6428,
; TITLE OF INVENTION: 7181, 7660, 25641, 69583, 49863, 8897, 1682, 17667, 9235,
; TITLE OF INVENTION: 3703, 14171, 10359, 1660, 1450, 18894, 2088, 32427, 2160,
; TITLE OF INVENTION: 9252, 9389, 1542, 85269, 10297, 1584, 9525, 14124, 4469,
; TITLE OF INVENTION: 8990, 2100, 9288, 64698, 10480, 20893, 33230, 1386, 9943,
; TITLE OF INVENTION: 16334, 68862, 9011, 14031, 6178, 21225, 1420, 32236, 2099,
; TITLE OF INVENTION: 2150, 26583, 2784, 8941, 9811, 27444, 50566 OR 66428 MOLECULES
; FILE REFERENCE: MP102-020P1RNMNM
; CURRENT APPLICATION NUMBER: US/10/354,358
; PRIOR FILING DATE: 2003-01-30
; PRIOR APPLICATION NUMBER: US 60/353,600
; PRIOR FILING DATE: 2002-01-31
; PRIOR APPLICATION NUMBER: US 60/364,517
; PRIOR FILING DATE: 2002-03-15
; PRIOR APPLICATION NUMBER: US 60/371,075
; PRIOR FILING DATE: 2002-04-09
; PRIOR APPLICATION NUMBER: US 60/371,507
; PRIOR FILING DATE: 2002-04-10
; PRIOR APPLICATION NUMBER: US 60/372,984
; PRIOR FILING DATE: 2002-04-16
; PRIOR APPLICATION NUMBER: US 60/374,194
; PRIOR FILING DATE: 2002-04-19
; PRIOR APPLICATION NUMBER: US 60/382,995
; PRIOR FILING DATE: 2002-05-24
; PRIOR APPLICATION NUMBER: US 60/385,023
; PRIOR FILING DATE: 2002-05-31
; PRIOR APPLICATION NUMBER: US 60/388,853
; PRIOR FILING DATE: 2002-06-14
; PRIOR APPLICATION NUMBER: US 60/389,395
; PRIOR FILING DATE: 2002-06-17
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 122
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 91
; LENGTH: 2980
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (86) ... (2422)
US-10-354-358-91

Query Match 46.6%; Score 851.4; DB 12; Length 2980;
Best Local Similarity 68.9%; Pred. No. 5.5e-224;
Matches 1356; Conservative 0; Mismatches 416; Indels 195; Gaps 6;

Qy 1 ATGAAGCTGGAGCGGAGATCGGATCCTGAGCTCATTGAGCACCCCGCTTAAG 60
Db 311 ATGAAGCTGGAGCGGAGATCGGATCCTGAGCTCATTGAGCACCCCGCTTAAG 370
Qy 61 CTGACGACGCTTTATGAAACCAAAAAATATTTGACCTGGTGTAGAACACGTCAGGT 120
Db 371 CTCAACGCTCTACGAGACAGAAATATTTGACCTGGTGTCTGGAGCAGCTCTCGGG 430
Qy 121 GGTGAGCTCTTCGACTACTCTGGTGAAGAGGGAGGTGACGCTTAAGAGGCTCGGAAG 180
Db 431 GGTGAGCTATTTCGACTACTCTGGTGAAGAGGGAGAGCTGACGCCCAAGAGGCGCGGAAG 490
Qy 181 TTCTTCGGGAGATCATCTCTCGCTGGACTTCTGCCACAGCCACTCCATATGCCACAGG 240
Db 491 TTCTTCGGGAGATTTGTTCTGCGCTGGACTTCTGCCACAGCTACTCCATCTGCCACAGA 550
Qy 241 GATCTGAAACCTGAAACCTCTCTCTGAGAGAGAACCAACATCCGCTCGCAGACTTT 300
Db 551 GACCTAAGCCGAGAACCTCTCTTTGGATGAGAAAAACAACATCCGCTATGCGACTTC 610
Qy 301 GGCATGGCGTCCCTGCGAGGTGGCGACAGCTGTGTGAGACCAAGCTGTGGGTCCCCCAC 360

```

Db 611 GGCAATGGCGTCCCTGAGGTGGGGAGAGCTCTCTGGAGACAGCTGCGGGTCCCGCCCAT 670
Qy 361 TAGCCTGCCCCAGGTGATCCGGGGGAGAGTATGACGGCGGGAAGGGGACGTGTGG 420
Db 671 TATGCGGTGTCAGAGGTGATTAAGGGGGGAAAAATATGATGGCGCGGGGAGACATGTGG 730
Qy 421 AGCTGGGGGCTCATCTCTGCTGCTGCTGGGGGCTCTGCGCTTTCAGCATGACAAAC 480
Db 731 AGCTGGAGTCACTCTCTGCTGCTGCTGGGGCTCTGCGCTTTCAGCATGACAAAC 790
Qy 481 TTGCGACAGTCTGAGAGAGGTGAAGCGGGGCTGTTCACATGCGCGCATTTATCCCG 540
Db 791 CTCGGCAGCTGTGAGAGAGGTGAAGCGGGGCTCTTCACATGCGCGCATTTATCCCT 850
Qy 541 CCGAGCTGCGAGAGTCTGCTACGGGGCATGATCGAGGTGAGCGCCGACCGCGCTCACG 600
Db 851 CCAGATTGCGAGAGCTCTGAGGGGGAATGATCGAAGTGGAGCCCGAAGAGGCTCAGT 910
Qy 601 CTAGAGCACATTTCAGAAACACATATGATATAGGGGGCAAGATGAGCCCGAACC- 656
Db 911 CTGAGGCAATTCAGAAACATCTTGTACTAGCGGGGAAACAGAGCCAGACCGCTGC 970
Qy 657 --AGAGCAGCCCATCTCTCGAAGGTGACATCGCTGCTGCTCCAGCTTGGAGGACATC 714
Db 971 CTGAGGACAGCCCTGCGCCCGGGGTAGCCATGCGAGCTGCGATCCCAACGGAGAGCTG 1030
Qy 715 GACCCGACGTGCGAGACATGCACTCACTGGGCTGCTTCGAGACCGCAACAGCTG 774
Db 1031 GACCCGACGTCTAGAGAGCATGCGCATCTAGCGCTGCTTCAGGACCGCGAGAGCTG 1090
Qy 775 CTGAGGACCTGCTGCTCGAGGAGGAGAACAGAGAGATGATTTACTTCTCTCTCTG 834
Db 1091 CATCGAGCTGCGCAGTGGAGGAGAGAACCAAGAAAGATGATATATCTGCTTTG 1150
Qy 835 GACCGAAAGAGGTATCCGAGCAGAGAGATGAGGACCTGCGCCCGCGGAAAGAGATA 894
Db 1151 GATCGAAGGAGCGGTATCCAGCTGTGAGGACAGGACCTGCGCTCCCGGGAATGATTT 1210
Qy 895 GACCTCTCCCGAGCGTGTGACTCCCGATGCTGAACCGGACGCGCAAGCGCGGCCA 954
Db 1211 GACCCCGCGGAGCGTGTGGATTCTCCATGCTGAGCGTCAACGGGAAGCGGGACCA 1270
Qy 955 GAACGCAAAATTCATGAGGTGCTCAGCGTGACGGA- 1005
Db 1271 GAGCGGAAGTCCATGGAAGTCTGAGCATCACGATCCCGGGGTGTGCTCCCTGTGA 1330
Qy 1006 CTTGCGGGCGGGCCATTGAGATGCGCCAGCGCCAGAGGTCTCGGTTCATCAGCGGT 1065
Db 1331 CCCACCGAGCGGGCTTGGAGATGGCCCGACAGCCAGAGATCCCGTAGCGTCAGTGGA 1390
Qy 1066 GCCTCTCAGGCTTTCCACGAGCCCACTCAGCAGCCCGCGG- 1108
Db 1391 GCCTTCAGGGTCTGTCTCCAGCCCTCTAAGCAGCCCAAGGAGTCCGGTCTTTCTTT 1450
Qy 1109 - 1108
Db 1451 TCACCGGAGCGGGGCTGGAGATGAGCTCGAGGCGGGGCTCCCGACTTCACAAAG 1510
Qy 1109 - 1122
Db 1511 CAGACGTGCTTCTCGGGGCCCCAGGGGTGGGGCGCGCGGAGAGCAGCCCGCGCCCC 1570
Qy 1123 TCACCAAGGGGCACTGCTCCCTCCCC- 1149
Db 1571 AGTCCGCTCCACACCCCTGCGGCCCCCAGGCTCCCGCGCTCTCTGCGGGGACC 1630
Qy 1150 CCCAAGGGGACACTGTCTCACAACCGCAAGAGAGCGCGGTGCGACGCCCAACCCCAAG 1209
Db 1631 CCTTTGCACTCGCTCTGCAACACCGCCCGGCGCAGTCCCAACCGGGAACCCCGGGGACA 1690
Qy 1210 CCCCCGTCAGCCCC- 1263
Db 1691 CCACCCCGAGCCCGCGGTGGCGCTCGGGGGAGCCCGCTGGAGGAGTCTGCTCAACTCC 1750

Qy 1264 ATCAGAAACAGCTTTCTGSGCTCACCCGCTTCCACCGCGGAAACTCAAGTTTCCGACG 1323
Db 1751 ATCCGCAACAGCTTCTGGGCTCCCGCTTCCAGCGCGCAAGATGAGGTCTCTTACC 1810
Qy 1324 CCGGAGGAGATGTCAACCTGACACACAGAGTGTGTCCTCCAGAGTGGCGAAGTCTCTGG 1383
Db 1811 GCTGAGGAGATGTCCAGCTTGACGCGCAGAGTCTCTCCCGAGCTGGCAAAACGCTCTGG 1870
Qy 1384 TTTGGGAATTCATCAGCTTGGAGAGGAGGAGAGATCTTCGTGTCATCAAAACACAAA 1443
Db 1871 TTCGGGAATTCATCTCTTGGACAAAGAAACAAATATTCCTGTCGTAAGGACAAA 1930
Qy 1444 CTTCTGAGCTCATCAAGGCTGACATCGTGACGCTTCTTCTGTCGATTCCTCAGC 1503
Db 1931 CTTCTCAGCAGATCAAGCAGACATCGTCCATGCTTCTGTCGATCCCCAGCTGAGT 1990
Qy 1504 CACAGCGTATCTCCAAACGAGCTTCCGGGCGGAGTACAGGCCACCGGGGGGCCAGCC 1563
Db 1991 CACAGTGTGCTGTCACAGACAGCTTTCAGGGCGGAGTACAAAGCCAGTGGCGGCCCTCC 2050
Qy 1564 GTGTTCCAGAAACCGGTCAGTTCCAGGTTGATATCACTACACGAGGGT- 1614
Db 2051 GTCTTCCAAAGCCCGTCGCTTTCAGGTTGAGCATCAGCTCTCTGAGGGTCCAGAGCCC 2110
Qy 1615 - 1665
Db 2111 TCCCGCGAGCGGACGGCAGCGAGGTGGTGGCATCTACTCCGTCACCTTCACTCTCATC 2170
Qy 1666 TCAGGCCCCAGCGCTGCTTCAAGAGGGTGGGAGACCATCCAGGCCAGCTGTGAGC 1725
Db 2171 TCGGGTCCAGCGCTGCTTCAAGCGAGTGGTGGAGACCATCCAGGCCAGCTCTGAGC 2230
Qy 1726 ACACAGCAGCCCGTCCGCGCCAGCACTTGTTCAGACACCACTAACTG 1772
Db 2231 ACTCATGACACGCCCTCCGTGCGAGCCCTGGCAGACGAGAAACGG 2277

RESULT 15

US-10-116-326-3

; Sequence 3, Application US/10116326

; Publication No. US20030166889A1

; GENERAL INFORMATION:

; APPLICANT: Turner, C. Alexander Jr.

; APPLICANT: Mathur, Brian

; APPLICANT: Fridge, Carl Johan

; TITLE OF INVENTION: No. US20030166889A1 Human Kinases and Polynucleotides Encoding t

; FILE REFERENCE: LEX-0332-USA

; CURRENT APPLICATION NUMBER: US/10/116,326

; PRIOR FILING DATE: 2002-04-04

; PRIOR FILING DATE: 2001-04-06

; NUMBER OF SEQ ID NOS: 6

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 3

; LENGTH: 2289

; TYPE: DNA

; ORGANISM: homo sapiens

US-10-116-326-3

Query Match

Best Local Similarity 42.5%; Score 777.2; DB 12; Length 2289;

Matches 1277; Conservative 67.9%; Pred. No. 1.4e-203;

Indels 195; Gaps 6;

Qy 88 TATTGTACTGCTGTAGAACACGCTGTGAGGTGGTGTGAGCTCTTCAGCTACCTGTTGTAAG 147
Db 265 TTTAGGTACTGCTGTCTGGAGCAGCTCTCGGGGGTGTGAGCTATTCAGCTACCTGTTAAAG 324
Qy 148 AAGGGAGGCTGTACGCTTAAGGAGGCTCGGAAGTTCTTCGGGCGAGATCATCTTCGGCTG 207
Db 325 AAGGGAGGACTGACGCGCCAGGAGGCGCCGGAAGTTCTTCGGCCAGATTGTGTGCGCTG 384
Qy 208 GACTTCTGCGACAGCCCACTCCATATATGCCACAGGGATCTGAAACCTGAAACCTCTCTGCTG 267

Db 385 |||||GACTTCTGCCACAGCTACTCCATCTGCCACAGAGACCTTAAGCCCGAGAACTGCTTTTG 444
Qy 268 GACGAGAAGAACAAATCCGATCCGATCCGAGACTTTGGCATGGGTCCCTGCGAGTTGGCGAC 327
Db 445 GATGAGAAACAAACATCCGATTTGCAGACTTCCGATGGGTCCCTGCGAGTTGGGGGAC 504
Qy 328 AGCTGTGGAGACCAAGTGTGGTTCGCCCACTACGCTGCCCGCCGAGGTGATCCGGGGG 387
Db 505 AGCTCTCTGGAGACCAAGTGTGGTTCGCCCACTACGCTGCCCGCCGAGGTGATTAAGGGG 564
Qy 388 GAGAAGTATGACGGCGGAGCGGACGTTGGAGCTTGGCGGTATCTGTTGCGCTTG 447
Db 565 GAAATAATGATGACGGCGGAGACATGTGGAGCTGTGGAGTCACTCTCTTCCGCCCTG 624
Qy 448 CTGTTGGGGGCTCTGCCCTTCCAGATGACAACTTGGACAGCTGCTGGAGAAAGTGAAG 507
Db 625 CTGTTGGGGGCTCTGCCCTTGGATGACGACAACTTCCGCCAGCTGCTGGAGAGTGAAG 684
Qy 508 CGGGCGGTGTTCCATGACCGCACTTTATCCGCCCGCACTGCCAGAGTCTGCTACGGGGC 567
Db 685 CGGGCGGTCTTCCATGACCGCACTTTATCCGCCAGCTTCCAGATTGCCAGAGCTCTGAGGGGA 744
Qy 568 ATGATCGAGTGGACCGCGAGCGGCTCAGCTAGAGACATTCAGAAACATATGG 627
Db 745 ATGATCGAAGTGGAGCGCGGAGGAGGCTCAGTCTGGAGCAAAATTCAGAAACATCTTGG 804
Qy 628 TATATAGGGGCAAGAAATGAGCCCGAACCC-----AGAGCAGCCCATTTCTCGCAAGGTG 681
Db 805 TACTAGCGGGAAACACAGACGACGACCCGCTGCTGGAGCCAGCCCTTGGCGCGCGGGTA 864
Qy 682 CAGATCCGCTCGCTGCCAGCTGGAGACATCGACCCCGAGCTGCTGGACAGCATGCAC 741
Db 865 GCCATCGGAGCTTGCATCCAAACGAGAGCTGGACCCCGAGCTCTAGAGAGCATGSCA 924
Qy 742 TCATCGGCTGCTTCCGAGCGGCAACAGCTGCTGAGGACCTGCTGTCGAGAGGAG 801
Db 925 TCATCGGCTGCTTCCGAGCGGAGGAGGCTGATCGCGAGCTGCGCAGTGGAGGAG 984
Qy 802 AACCAGGAGAGATGATTTACTTCTCTCTCGACCGGAAAGAAAGTACCCGAGCCAG 861
Db 985 AACCAGGAGAGATGATATATTTCTGTTTGGATCGGAGAGCGGTATCCAGCTGT 1044
Qy 862 GAGGATGAGGACCTGCCCCCGGAAACAGATAGACCTTCCCGGAAAGCGTGTGACTCC 921
Db 1045 GAGGACCAAGGACCTGCTTCCCGGAAATGATGTTGACCCCGGAGCGGTGGAATCT 1104
Qy 922 CCGATGCTGACCGGACGAGCGGCGGCGGACGACGCAATCCATGGAGTCTCAGC 981
Db 1105 CCGATGCTGAGCGGTCACGGGAAGCGGCGGACGAGAGCGGAAATCCATGGAAGTCTTGAGC 1164
Qy 982 GTGAC-----GGACGGCGGCTCCCGGTGCTGCGCGCGGCGCATTTGAGATGGCC 1032
Db 1165 ATCACCAGTACCGGGGGTGGTGGCTCCCTGTATCCACCGACGGGCTTGGAGATGGCC 1224
Qy 1033 CAGCAGCGCCAGAGGTCTCGGTTCATAGCGGTGCTCTCTCAGSCCTTTCCACAGCCCA 1092
Db 1225 CAGCAGCCAGAGATCCCGTAGCGTCACTGAGGAGCCTCCAGGGTCTGTCTCTCAGCCCT 1284
Qy 1093 CTCAGCAGCCCG----- 1106
Db 1285 CTAAGCAGCCCAAGAGTCCGGTCTTTTCTTTTCAACGAGCGCGGGGTGGAGATGAG 1344
Qy 1107 ----- 1106
Db 1345 GCTCGAGCGGGGTCCTCCGACTTCCAAAACGAGAGCTGCTTCTCGGGGCCCCAGG 1404
Qy 1107 -----GGTACCCCTTCAACCCCTCAACAAAGGGGAGTCCCTTCCCC--- 1146
Db 1405 GGTGGGGCGCGGGGAGCAGCCCGCGCCCGAGTGCCTCCACACCCCTTCCCGGC 1464
Qy 1147 -----ACCCCAAGGGGACACCTGTTCACAGCCCA 1176
|||||

Db 1465 CCCCCAGGCTCCCGCGCTCTCTGTGGGGGACCCCTTGCACTCGCTCTGCAACAGCCC 1524
Qy 1177 AAGGAGAGCCCGGCTGCGACGCGCCAAACCCAGCCCGCTCCAGCCCC-----AGCGTC 1230
Db 1525 CGGGCCAGTCCACACCGGACCCCGGGGACAAACACCCCGGCGCGGTGGCGTC 1584
Qy 1231 GGAGGGTCCCTGGAGGCGGGCTCAACTCCATCAAGAACAGCTTTCTGGGCTCACCC 1290
Db 1585 GGGGAGCGCTGAGAGTCTCTCAACTCCATCCGCAACAGCTTCTTGGGCTCCCT 1644
Qy 1291 CGCTTCCACCGCCCGAAATCTGAAAGTTCGACGCCCGGAGAGATGTCAACTGACCA 1350
Db 1645 CGCTTTCACCGGCGCAAGATGCAAGTCCCTTACCGCTGAGGAGATGTCACGCTTGA 1704
Qy 1351 GAGTCTGCTCCCGAGCTGCGAAGAGTCTGTGTTGGGAATTTCAATCAGCCTGGAGAAG 1410
Db 1705 GAGTCTCTCCCGAGCTGCGAAGAGTCTGTGTTGGGAACTTCACTCTCTTGGACAA 1764
Qy 1411 GAGGAGCAGATCTCTGTGTCATCAAAAGACAAACCTCTGAGCTCCATCAAGGCTGACATC 1470
Db 1765 GAAGAACAAATATTCCTGCTGCTAAAGGACAAACCTCTCAGCAGCATCAAAGCAGCATC 1824
Qy 1471 GTCCAGCGCTTCTGTGTCATCCAGTCTCAGCAGCAGGTCATCTCCAAACAGGCTTC 1530
Db 1825 GTCCATGCTTCTGTGTCATCCCGAGCTCCCGCTCAGTGTGCTGTACAGACCGCTTC 1884
Qy 1531 CGGGCCGAGTACAAAGCCACCGGGGGCCAGCGGTGTCAGAAAGCGGTCAAGTTCCAG 1590
Db 1885 AGGGCCGAGTACAAAGCCAGTGTGCGGCCCTCCCTCTTCCAAAGCCCGTCCCGCTTCAG 1944
Qy 1591 GTTGATATCACTTACACGGAGGT-----GGGAGGCGCAGAAAGGAG 1632
Db 1945 GTGACATCAGCTCTCTGAGGCTCCAGAGCCCTCCCGCGACGCGACGCGCAGCGAGGT 2004
Qy 1633 AAGGATCTACTCCCTGCTTCACTGCTCTCAGGCCCCAGCGCTTCAAGAGG 1692
Db 2005 GGTGGCATCTACTCCGCTCACTCTCTCATCTCGGGTCCCGAGCGTCCGTTCAAGCGA 2064
Qy 1693 GTGGTGGAGACCATCCAGGCCAGCTGTGAGCACACAGCCCGCTGCGGCCAGCAC 1752
Db 2065 GTGGTGGAGNCCATCCAGGCGACAGCTCTCTGAGCACTCATGACCAGCCCTCCGTGCA 2124
Qy 1753 TTGTCAGACACCACTAACTG 1772
Db 2125 CTGGCAGACGAGAAGAACGG 2144

Search completed: November 27, 2003, 15:02:51
Job time : 1072.75 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 27, 2003, 01:57:44 ; Search time 501.488 Seconds
(without alignments)
10803.388 Million cell updates

Title: US-10-054-579-1

Perfect score: 2007

Sequence: 1 atgacatcgacggggaagga.....gaattatccgaaaagttaa 2007

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 2552756 seqs, 1349719017 residues

Total number of hits satisfying chosen parameters: 5105512

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum March 0%

100% Processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database : N Geneseq 19Jun03:*

1:	/SDSI1/cgcgdata/geneseq/geneseqn-emb1/NA1980.DAT.*
2:	/SDSI1/cgcgdata/geneseq/geneseqn-emb1/NA1981.DAT.*
3:	/SDSI1/cgcgdata/geneseq/geneseqn-emb1/NA1982.DAT.*
4:	/SDSI1/cgcgdata/geneseq/geneseqn-emb1/NA1983.DAT.*
5:	/SDSI1/cgcgdata/geneseq/geneseqn-emb1/NA1984.DAT.*
6:	/SDSI1/cgcgdata/geneseq/geneseqn-emb1/NA1985.DAT.*
7:	/SDSI1/cgcgdata/geneseq/geneseqn-emb1/NA1986.DAT.*
8:	/SDSI1/cgcgdata/geneseq/geneseqn-emb1/NA1987.DAT.*
9:	/SDSI1/cgcgdata/geneseq/geneseqn-emb1/NA1988.DAT.*
10:	/SDSI1/cgcgdata/geneseq/geneseqn-emb1/NA1989.DAT.*
11:	/SDSI1/cgcgdata/geneseq/geneseqn-emb1/NA1990.DAT.*
12:	/SDSI1/cgcgdata/geneseq/geneseqn-emb1/NA1991.DAT.*
13:	/SDSI1/cgcgdata/geneseq/geneseqn-emb1/NA1992.DAT.*
14:	/SDSI1/cgcgdata/geneseq/geneseqn-emb1/NA1993.DAT.*
15:	/SDSI1/cgcgdata/geneseq/geneseqn-emb1/NA1994.DAT.*
16:	/SDSI1/cgcgdata/geneseq/geneseqn-emb1/NA1995.DAT.*
17:	/SDSI1/cgcgdata/geneseq/geneseqn-emb1/NA1996.DAT.*
18:	/SDSI1/cgcgdata/geneseq/geneseqn-emb1/NA1997.DAT.*
19:	/SDSI1/cgcgdata/geneseq/geneseqn-emb1/NA1998.DAT.*
20:	/SDSI1/cgcgdata/geneseq/geneseqn-emb1/NA1999.DAT.*
21:	/SDSI1/cgcgdata/geneseq/geneseqn-emb1/NA2000.DAT.*
22:	/SDSI1/cgcgdata/geneseq/geneseqn-emb1/NA2001A.DAT
23:	/SDSI1/cgcgdata/geneseq/geneseqn-emb1/NA2001B.DAT
24:	/SDSI1/cgcgdata/geneseq/geneseqn-emb1/NA2002.DAT.*
25:	/SDSI1/cgcgdata/geneseq/geneseqn-emb1/NA2003.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	2007	100.0	2007	24	ABQ77625	Human kinase NHP #
2	2005.4	99.9	2908	25	ABX13583	Human cDNA encodin
3	1993	99.3	3364	25	ABX13584	Human cDNA encodin
4	1966	98.0	2647	24	AAD34315	Human PRIN-18 cDNA
5	1939.6	96.6	2025	24	ABA02995	Human protein kina
6	1939.6	96.6	2217	24	ABA02994	Human protein kina
7	1827	91.0	1827	24	ABQ77626	Human kinase NHP #
8	1017.8	50.7	2799	25	AAL51889	Human cell cycle-r

9	965	48.1	2337	25	ABV74557	Human kinase #1 co
10	964.4	48.1	2385	22	AAS06717	Polynucleotide seq
11	964.4	48.1	2897	24	AAD26454	Human kinase PKIN
12	851.4	42.4	2112	25	ABV74559	Human kinase #3 co
13	777.2	38.7	2289	25	ABV74558	Human kinase #2 co
14	588.8	29.3	1797	25	ABX70994	Novel human cDNA s
15	539.8	26.9	2720	23	ABL10489	Drosophila melanog
16	449.2	22.4	906	22	ABA08296	Human HrPOK-1 hom
17	291.6	14.5	512	24	ABK70216	Human lung cancer
18	265.8	13.2	3609	25	AAD51408	Human microtubule
19	265	13.2	1594	22	AAFA4655	Novel protein kina
20	264.2	13.2	1549	24	ABS64386	Human serine/threo
21	264.2	13.2	2085	25	AAD51409	Human microtubule
22	264.2	13.2	2278	25	AAD51410	Human microtubule
23	264.2	13.2	2462	24	ABQ93445	Human cDNA SEQ ID
24	264.2	13.2	2954	22	AAK94923	Human full-length
25	264.2	13.2	3226	25	AAD51412	Human microtubule
26	264.2	13.2	3269	22	ADAD18831	Human kinase (PKIN
27	264.2	13.2	3312	24	ABA05739	Human neuronal ser
28	264.2	13.2	3392	24	ABA05740	Human neuronal ser
29	264.2	13.2	4917	25	AAD51411	Human microtubule
30	262.6	13.1	3270	25	ABQ77242	Human MARK cDNA
31	262.4	13.1	3170	24	ABA05737	Murine neuronal se
32	262.4	13.1	3250	24	ABA05738	Murine neuronal se
33	261.4	13.0	2663	23	ABL05573	Drosophila melanog
34	252.4	12.6	2352	22	AAI16683	Human protein kina
35	252.4	12.6	2968	22	AAI66822	Human protein kina
36	252.4	12.6	4699	22	AAH76213	Human kinase PKIN
37	250.8	12.5	2632	21	AAK82951	Human keratinocyte
38	250.8	12.5	2632	21	AAK82952	Human keratinocyte
39	249.8	12.4	614	22	ABA08924	Human Ser/Thr kina
40	233.4	11.6	2118	23	ABL11605	Drosophila melanog
41	233.4	11.6	2361	22	AAS06718	Polynucleotide seq
42	233.4	11.6	2361	23	ABL11604	Drosophila melanog
43	232.6	11.6	2766	23	ABL26615	Drosophila melanog
44	230.6	11.5	3234	25	ABZ77147	Human protein kina
45	229.4	11.4	2601	23	ABL13339	Drosophila melanog

ALIGNMENTS

RESULT 1	
ABQ07625	
ID	ABQ077625 standard; cDNA; 2007 BP.
XX	
XX	ABQ077625;
DT	21-OCT-2002 (first entry)
XX	
DE	Human kinase NHP #1-encoding cDNA, SEQ ID NO:1.
XX	
KW	Human; Kinase; NHP; chromosome 11; signal transduction; nuclear; brain; pituitary; hypothalamus; adipose tissue; adrenal gland; fetal lung; fetal brain; cancer; mental disorder; diagnostic reagent; clinical trial; nutraceutical; mutation detection; gene expression; chromosome mapping; transgenic animal; cytostatic gene; ss.
XX	
OS	Homo sapiens.
XX	
FH	Key Location/Qualifiers
CDS	1..2007
FT	/*tag= a
FT	/product= "Human kinase NHP #1"
FT	
PN	WO200259287-A2.
XX	
PD	01-AUG-2002.
XX	
PF	22-JAN-2002; 2002WO-US01818.
XX	

23-JAN-2001; 2001US-263378P.
(LEXI-) LEXICON GENETICS INC.
Turner CA, Mathur B;
WPI; 2002-599780/64.
P-PSDB; ABB09557.
Novel polynucleotide encoding human proteins sharing sequence
similarity with animal kinases, useful for drug screening, diagnosis,
in gene therapy of disorders and diseases e.g. cancer -
Claim 1; Page 36; 40pp; English.
The invention relates to novel human kinases (ABB09557, AAB09558),
referred to as NHPs, and to nucleic acids encoding them (ABQ77625,
ASQ77626). The NHPs have structural similarity with animal kinases
including serine-threonine kinases, carbon catabolite repressing kinases,
calcium/calmodulin-dependent protein kinases, and cAMP-dependent protein
kinases. The NHPs do not contain consensus signal sequences, indicating
that they may be cytoplasmic or nuclear proteins, and are thought to be
involved in signal transduction. Polynucleotides encoding NHPs were
obtained using human brain, pituitary, hypothalamus, adipose, cerebellum,
adrenal gland, foetal lung, and foetal brain cDNAs and primers derived
from human genomic DNA. The gene encoding the NHPs is located on
chromosome 11. NHP nucleotides and proteins are useful for treating
mental, biological or medical disorders including cancer, and for
screening compounds useful in the treatment of such conditions. They can
also be used as diagnostic reagents, in clinical trial monitoring and in
cosmetic and nutritional applications. NHP nucleotides can additionally
be used in the detection of disease-associated mutations, in the analysis
of gene expression, for mapping chromosome 11, for the recombinant
expression of NHPs, to generate transgenic animals, in gene therapy,
and as part of ribozyme and/or triple helix sequences useful in the
modulation of NHP gene expression. The present sequence represents cDNA
encoding a 668 amino acid NHP (#1).
Sequence 2007 BP; 445 A; 626 C; 604 G; 332 T; 0 other;

Query Match 100.0%; Score 2007; DB 24; Length 2007;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2007; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 ATGACATCGACGGGGAAGGACGGCGGCGCAGCACGCGAGTATGTTGGGCCCTACCGG 60
1 ATGACATCGACGGGGAAGGACGGCGGCGCAGCACGCGAGTATGTTGGGCCCTACCGG 60
61 CTGAGAAGACGCTGGGCAAGGGGCGACAGAGTCTGTGTGAAGCTGGGGGTTCACTGGCTC 120
61 CTGAGAAGACGCTGGGCAAGGGGCGACAGAGTCTGTGTGAAGCTGGGGGTTCACTGGCTC 120
121 ACCTGCCAGAGGTGGCCATCAAGATCGTCAACCGTGCAGAGCTCAGCGAGTGGTCTG 180
121 ACCTGCCAGAGGTGGCCATCAAGATCGTCAACCGTGCAGAGCTCAGCGAGTGGTCTG 180
181 ATGAAGGTGGAGCGGGAGATCGCATCTGAAGTCTATTGAGCACCCCGCAGTCTTAAG 240
181 ATGAAGGTGGAGCGGGAGATCGCATCTGAAGTCTATTGAGCACCCCGCAGTCTTAAG 240
241 CTGACAGACGTTTATGAAAAAATAATTTGTTACTGTGTGTAGAACACGTTGTAGGT 300
241 CTGACAGACGTTTATGAAAAAATAATTTGTTACTGTGTGTAGAACACGTTGTAGGT 300
301 GGTGAGCTCTTCGACTACCTGGTGAAGGGGAGGCTGAGCGCTTAGAGGGCTCGGAAG 360
301 GGTGAGCTCTTCGACTACCTGGTGAAGGGGAGGCTGAGCGCTTAGAGGGCTCGGAAG 360
361 TTCTTCGGCAGATCATCTCTGCGCTGAGACTTCTGCCACAGCCACTCCATATGCCACAGG 420
361 TTCTTCGGCAGATCATCTCTGCGCTGAGACTTCTGCCACAGCCACTCCATATGCCACAGG 420
421 GATCTGAAAACCTGAAAACCTCTCTGCTGGACGAGAAGAACAAACATCCGCATCGCAGCTTT 480

421 GATCTGAAAACCTGAAAACCTCTCTGCTGGACGAGAAGAACAAACATCCGCATCGCAGCTTT 480
481 GGCATGGCGTCCCTGCAGGTTGGCGACAGCCCTGTTGGAGACCCAGCTGTGGTCCCGCCAC 540
481 GGCATGGCGTCCCTGCAGGTTGGCGACAGCCCTGTTGGAGACCCAGCTGTGGTCCCGCCAC 540
541 TACGCTCTGCCCGAGGTGATCCGGGGGGAAGATATGACGGCGGAAGCGGACGCTGTGG 600
541 TACGCTCTGCCCGAGGTGATCCGGGGGGAAGATATGACGGCGGAAGCGGACGCTGTGG 600
601 AGCTCGGCGTCACTCTGTTGCGCTTGTGTGGGGGCTCTGCCCTTGCACGATGACAAAC 660
601 AGCTCGGCGTCACTCTGTTGCGCTTGTGTGGGGGCTCTGCCCTTGCACGATGACAAAC 660
661 TTGCCACAGCTGCTGAGAGGTGAAGGGGCGGTTCACATCGCGCACCTTTATCCCG 720
661 TTGCCACAGCTGCTGAGAGGTGAAGGGGCGGTTCACATCGCGCACCTTTATCCCG 720
721 CCCGACTGCGAGAGTCTGCTACGGGGGCATGATCGAGGTGGACGCGCGCTCTCACG 780
721 CCCGACTGCGAGAGTCTGCTACGGGGGCATGATCGAGGTGGACGCGCGCTCTCACG 780
781 CTAGAGCATTTTCAGAAAACATATGTTATAGGGGGCAAGATAGCCCCGAAACAGAG 840
781 CTAGAGCATTTTCAGAAAACATATGTTATAGGGGGCAAGATAGCCCCGAAACAGAG 840
841 CAGCCCATCTTCGCAAGGTGCAGATCCGCTCGCTGCCAGCTCGAGGACATCGACCCC 900
841 CAGCCCATCTTCGCAAGGTGCAGATCCGCTCGCTGCCAGCTCGAGGACATCGACCCC 900
901 GACGTGCTGGACAGCATGCACTCACTGGGCTGCTTCCGAGACCGCAACAGCTGTGCAG 960
901 GACGTGCTGGACAGCATGCACTCACTGGGCTGCTTCCGAGACCGCAACAGCTGTGCAG 960
961 GACCTGCTCTCGAGGAGAGAACAGGAGAGATGATTTACTTCTCTCTCTGACCGG 1020
961 GACCTGCTCTCGAGGAGAGAACAGGAGAGATGATTTACTTCTCTCTCTGACCGG 1020
1021 AAGAAGGTGTACCCGAGCAGGAGATGAGGACCTGCCCCCGGAAACAGATAGACCTT 1080
1021 AAGAAGGTGTACCCGAGCAGGAGATGAGGACCTGCCCCCGGAAACAGATAGACCTT 1080
1081 CCCGGAAGCTGTGGAATCCCCCGATGCTGAACCGGACGCGCAAGCGGCGGCAAGCGC 1140
1081 CCCGGAAGCTGTGGAATCCCCCGATGCTGAACCGGACGCGCAAGCGGCGGCAAGCGC 1140
1141 AAATCCATGGAGGTGCTCAGCTGACGCGGAGCTCCCGGTGCTGCTGCGGGCGGGCC 1200
1141 AAATCCATGGAGGTGCTCAGCTGACGCGGAGCTCCCGGTGCTGCTGCGGGCGGGCC 1200
1201 ATTGAGATGGCCAGCAGCGCCAGAGGTCTCGGTTCATCAGCGGTGCTCTCTCAGGCTT 1260
1201 ATTGAGATGGCCAGCAGCGCCAGAGGTCTCGGTTCATCAGCGGTGCTCTCTCAGGCTT 1260
1261 TCCACAGCGCTCTCAGAGCGCCCGGGTGAACCCCTCACCCCTCACAAAGGGGAGTCCC 1320
1261 TCCACAGCGCTCTCAGAGCGCCCGGGTGAACCCCTCACCCCTCACAAAGGGGAGTCCC 1320
1321 CTCCCCACCCCAAGGGGACACTGTTCCACAGCGCCCAAGAGAGCCCGGTGCGACGCC 1380
1321 CTCCCCACCCCAAGGGGACACTGTTCCACAGCGCCCAAGAGAGCCCGGTGCGACGCC 1380
1381 AACCCCAACCCCGCTTCACGCGCCAGCTCGGAGGGGTCCCTTGAAGGGCGCGGCTCAAC 1440
1381 AACCCCAACCCCGCTTCAGCGCCAGCTCGGAGGGGTCCCTTGAAGGGCGCGGCTCAAC 1440
1441 TCCATCAAGAAACAGCTTTCTGGGCTCAACCCCGTTCACCGCGGAAACTGCAAGTTCCG 1500
1441 TCCATCAAGAAACAGCTTTCTGGGCTCAACCCCGTTCACCGCGGAAACTGCAAGTTCCG 1500
1501 ACGCCGAGGAGATGTTCCAACTGACACAGAGTCTGTTCCCGCAGCTGCGGAGAGATCC 1560

Db 1501 AGCGCGAGGAGATCCAACTGACACAGAGTCGTCCCGAGAGCTGGCGAAGAGTCC 1560
Qy 1561 TGGTTTGGGAATCTCATCAGCGCTGAGAGAGGAGCAGATCTTCGTGGTTCATCAAGAC 1620
Db 1561 TGGTTTGGGAATCTCATCAGCGCTGAGAGAGGAGCAGATCTTCGTGGTTCATCAAGAC 1620
Qy 1621 AAACCTCTGAGTTCATCAGCGTGACATCTGTGACAGCGCTTCCTGTGATTCACAGTCTC 1680
Db 1621 AAACCTCTGAGTTCATCAGCGTGACATCTGTGACAGCGCTTCCTGTGATTCACAGTCTC 1680
Qy 1681 AGCCACAGCTCATCTCCCAACAGAGCTTCGGGCGGAGTCAAGGCCACCGGGGGGCCA 1740
Db 1681 AGCCACAGCTCATCTCCCAACAGAGCTTCGGGCGGAGTCAAGGCCACCGGGGGGCCA 1740
Qy 1741 GCCGTGTTCCAGAACCGGTCAAGTTCAGGTTGATATCACCTACACGAGGGTGGGGAG 1800
Db 1741 GCCGTGTTCCAGAACCGGTCAAGTTCAGGTTGATATCACCTACACGAGGGTGGGGAG 1800
Qy 1801 GCGGAGAGGAGAACGGATCTACTCCGTACCTTCACCTTCAGCGCTTCAGGCGCCAGCGGT 1860
Db 1801 GCGGAGAGGAGAACGGATCTACTCCGTACCTTCACCTTCAGCGCTTCAGGCGCCAGCGGT 1860
Qy 1861 CGCTTCAAGAGGGTGTGAGACCATCCAGGCCAGCTGTGAGCACACACGACCGCGCT 1920
Db 1861 CGCTTCAAGAGGGTGTGAGACCATCCAGGCCAGCTGTGAGCACACACGACCGCGCT 1920
Qy 1921 GCGGCCAGCACTTGTGACACACCACTAACTGTATGGAATGATGACGGGGCGCTTTCC 1980
Db 1921 GCGGCCAGCACTTGTGACACACCACTAACTGTATGGAATGATGACGGGGCGCTTTCC 1980
Qy 1981 AAATGTGAATTTATCCGAAAGTTAA 2007
Db 1981 AAATGTGAATTTATCCGAAAGTTAA 2007

RESULT 2

ABX13583
ID ABX13583 standard; cDNA; 2908 BP.
XX AC ABX13583;
XX AC
DT 10-FEB-2003 (first entry)
XX Human cDNA encoding serine protein kinase KSE336-1.
DE Human; ss; gene; chromosome 11p15.5-pter; astrocytoma; meningioma;
KW pancreatic adenocarcinoma; insulin-dependent diabetes mellitus 2;
KW helicoid peripapillary chorioretinal degeneration; brain; pancreas;
KW Beckwith-Wiedemann syndrome; congenital hyperinsulinism; KSE336.
XX Homo sapiens.
XX Key Location/Qualifiers
FT CDS 106..2112
FT /*tag= a
FT /product= "KSE336-1"
XX
XX US6455292-B1.
XX
XX 24-SEP-2002.
XX
XX 16-AUG-2001; 2001US-0930181.
XX
XX 16-AUG-2001; 2001US-0930181.
XX (ORIG-) ORIGENE TECHNOLOGIES INC.
XX
XX Shu Y, Fan W, Kovacs KF, Zidanic M, Jay G;
XX
XX WPI; 2003-066233/06.
XX P-PSDB; ABG72382.
XX
XX New isolated polynucleotide coding without interruption for a human

PT KSE336 polypeptide useful for preventing or treating
PT diseases/conditions relating to brain and pancreas, e.g. meningioma,
PT insulin-dependent diabetes mellitus 2
XX Claim 3; Column 33-40; 34pp; English.
XX
CC The invention relates to an isolated polynucleotide (its complement
CC or a sequence 99% similar to it) coding without interruption for a human
CC KSE336 polypeptide, a serine protein kinase, comprising the KSE336-1 and
CC KSE336-2 splice variants appearing as ABG72382 and ABG72383. Also
CC included is a method of identifying an agent that modulates the
CC expression of KSE336 in brain, pancreas, brain progenitor or pancreas
CC progenitor cells comprising: (a) contacting a cell population comprising
CC the cells with a test agent under conditions effective for the test agent
CC to modulate the expression of KSE336; and (b) determining if the test
CC agent modulates the expression of KSE336. The polynucleotides are useful
CC as molecular targets or drug targets, and for detecting, diagnosing, or
CC staging, monitoring, prognosticating, preventing or treating diseases or
CC conditions relating to brain and pancreas, such as astrocytoma,
CC meningioma, pancreatic adenocarcinoma, insulin-dependent diabetes
CC mellitus 2, helicoid peripapillary chorioretinal degeneration, Beckwith-
CC Wiedemann syndrome or congenital hyperinsulinism. The method and
CC polynucleotides are useful in research, diagnosis, drug discovery,
CC therapy, clinical medicine, forensic science and pathology.
CC The gene for KSE336 is located on chromosome 11p15.5-pter.
CC The present sequence encodes the KSE336 variant KSE366-1.
XX
SQ Sequence 2908 BP; 603 A; 948 C; 879 G; 478 T; 0 other;
Query Match 99.9%; Score 2005.4; DB 25; Length 2908;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2006; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
Qy 1 ATGACATCGACGGGGAAGGACGGCGCGCAGCACCGCAGTATGTGGGCCCTACCGG 60
Db 106 ATGACATCGACGGGGAAGGACGGCGCGCAGCACCGCAGTATGTGGGCCCTACCGG 165
Qy 61 CTGAGAGAAGCGCTGGGCNAAGGGGACAGACAGGTCTGTGAAGCTGGGGGTTCACTCGGTC 120
Db 166 CTGAGAGAAGCGCTGGGCNAAGGGGACAGACAGGTCTGTGAAGCTGGGGGTTCACTCGGTC 225
Qy 121 ACCTGCCAGAGGTGGGCCATCAGATCGTCACCGCTGAGAGCTCAGCGAGTCCGTGCTG 180
Db 226 ACCTGCCAGAGGTGGGCCATCAGATCGTCACCGCTGAGAGCTCAGCGAGTCCGTGCTG 285
Qy 181 ATGAAGGTGGAGCGGGAGATCGCGATCTGAAGCTCATTTAGACACCCCGCTCTAAAG 240
Db 286 ATGAAGGTGGAGCGGGAGATCGCGATCTGAAGCTCATTTAGACACCCCGCTCTAAAG 345
Qy 241 CTGACACGAGCTTTATGAAAAAATAATTTGTAACCTGTGTGTAGAACACGTCAGGT 300
Db 346 CTGACACGAGCTTTATGAAAAAATAATTTGTAACCTGTGTGTAGAACACGTCAGGT 405
Qy 301 GGTGAGCTCTTCGACTACCTGTTGAAGAAGGGGAGCTGACCGCTAAGGGGCTCGGAAG 360
Db 406 GGTGAGCTCTTCGACTACCTGTTGAAGAAGGGGAGCTGACCGCTAAGGGGCTCGGAAG 465
Qy 361 TTCTTCCGGCAGATCATCTCTGCGCTGAGCTTCTGCGCACAGCCACTCATATGCCACAGG 420
Db 466 TTCTTCCGGCAGATCATCTCTGCGCTGAGCTTCTGCGCACAGCCACTCATATGCCACAGG 525
Qy 421 GATCTGAAAACTGAAAACTCTCTGCTGGAGAGAGAACAAACATCCGCATCGCAGACTTT 480
Db 526 GATCTGAAAACTGAAAACTCTCTGCTGGAGAGAGAACAAACATCCGCATCGCAGACTTT 585
Qy 481 GGCATGGCGTCCCTGCGAGGTGGCGACAGCTTGTGGAGACAGCTGTGGTCCGCCAC 540
Db 586 GGCATGGCGTCCCTGCGAGGTGGCGACAGCTTGTGGAGACAGCTGTGGTCCGCCAC 645
Qy 541 TACGCTGCCCGGAGGTGATCCGGGGGAGAGATGATGACCGCGGAGGCGGAGCTGTGG 600
Db 646 TACGCTGCCCGGAGGTGATCCGGGGGAGAGATGATGACCGCGGAGGCGGAGCTGTGG 705

601 AGCTGGCGGCTCATCTCTGTTGCTGCTGGTGGGGCTCTGCCCTTTCGACGATGACAAC 660
Db AGCTGGCGGCTCATCTCTGTTGCTGCTGGTGGGGCTCTGCCCTTTCGACGATGACAAC 765
661 TTGGGACAGTCTGGAGAGGTGAAAGCGGGCGTGTTCACATGCCGCACTTTATCCCG 720
Db TTGGGACAGTCTGGAGAGGTGAAAGCGGGCGTGTTCACATGCCGCACTTTATCCCG 825
721 CCGGACTGCCAGAGTCTGCTTACCGGGGATGATCGAGGTGACCGCGCACGCCGCTCACG 780
Db CCGGACTGCCAGAGTCTGCTTACCGGGGATGATCGAGGTGACCGCGCACGCCGCTCACG 885
781 CTAGAGCACATTGAGAAACATATATGATATATAGGGGCAAGATGAGCCCGAACAGAG 840
Db CTAGAGCACATTGAGAAACATATATGATATATAGGGGCAAGATGAGCCCGAACAGAG 945
841 CAGCCATTCTCGCAGGTGCGATCCGCTCCGCTGCCAGCTGGAGGACATCGACCCC 900
Db CAGCCATTCTCGCAGGTGCGATCCGCTCCGCTGCCAGCTGGAGGACATCGACCCC 1005
901 GACGTGCTGGACAGCATGCACTCACTGGGCTGCTTCCGAGACCGCAACAAGCTGCTGCAG 960
Db GACGTGCTGGACAGCATGCACTCACTGGGCTGCTTCCGAGACCGCAACAAGCTGCTGCAG 1065
961 GACCTGCTCCGAGGAGAGAACACAGAGAAATGATTTACTTCTCTCTCGGACCGG 1020
Db GACCTGCTCCGAGGAGAGAACACAGAGAAATGATTTACTTCTCTCTCGGACCGG 1125
1021 AAGAAAGGTACCCGACGAGGAGTGAAGCTGCTGCCCGGACGAGATAGACCTT 1080
Db AAGAAAGGTACCCGACGAGGAGTGAAGCTGCTGCCCGGACGAGATAGACCTT 1185
1081 CCCCAGAGGTGTGAGTCTCCCGATCTGAACCGGACGCGCAAGCGCGCCAGAACGC 1140
Db CCCCAGAGGTGTGAGTCTCCCGATCTGAACCGGACGCGCAAGCGCGCCAGAACGC 1245
1141 AATTCATGAGGTGCTCAGGTGACGAGCGGCGTCCCGGCTGCTGCGCGCGGCGCC 1200
Db AATTCATGAGGTGCTCAGGTGACGAGCGGCGTCCCGGCTGCTGCGCGCGGCGCC 1305
1201 ATTGAGATGGCCAGCAGCGGCTCGGTCTCATCAGCGTGCCTCTCAGGCTT 1260
Db ATTGAGATGGCCAGCAGCGGCTCGGTCTCATCAGCGTGCCTCTCAGGCTT 1365
1261 TCCACACGCCCACTCAGCAGCCCCCGGTGACCCCTCACCCCTCACCAAGGGGCGTCCC 1320
Db TCCACACGCCCACTCAGCAGCCCCCGGTGACCCCTCACCCCTCACCAAGGGGCGTCCC 1425
1321 CTCCCGACCCCAAGGGGACACTGTTCACAGCGCAAGGAGAGCCCGCTGGACGCC 1380
Db CTCCCGACCCCAAGGGGACACTGTTCACAGCGCAAGGAGAGCCCGCTGGACGCC 1485
1381 AACCCACGCCCGCTCCAGCGCCAGCGTCCGAGGGGCTCCCTGAGGGGCGGCTCAAC 1440
Db AACCCACGCCCGCTCCAGCGCCAGCGTCCGAGGGGCTCCCTGAGGGGCGGCTCAAC 1545
1441 TCCATCAAGAACAGTCTTCTGGGCTCACCCGCTTCCACCGCCGGAACCTGCAAGTTCCG 1500
Db TCCATCAAGAACAGTCTTCTGGGCTCACCCGCTTCCACCGCCGGAACCTGCAAGTTCCG 1605
1501 ACGCCGAGGAGATGTCACCTGACACAGAGTGTCTCCAGAGCTGGCGAAGTCC 1560
Db ACGCCGAGGAGATGTCACCTGACACAGAGTGTCTCCAGAGCTGGCGAAGTCC 1665
1606 ACGCCGAGGAGATGTCACCTGACACAGAGTGTCTCCAGAGCTGGCGAAGTCC 1665
1661 TGGTTGGGAATTCATCAGCTGAGAGGAGGAGGAGTCTCGTGGTTCATCAAGAGAC 1620
Db TGGTTGGGAATTCATCAGCTGAGAGGAGGAGGAGTCTCGTGGTTCATCAAGAGAC 1725
1621 AAACCTCTGAGTCCATCAGGCTGACATCGTGACCGCTTCTGTGATTTCCAGTCTC 1680
Db AAACCTCTGAGTCCATCAGGCTGACATCGTGACCGCTTCTGTGATTTCCAGTCTC 1785
1681 AGCCACAGGTCTATCTCCCAACAGAGCTTCCGGGCGGAGTACAGGCCACCGGGGGGCA 1740

1786 AGCCACAGCGTCACTCTCCAAACGAGCTTCCGGGCCGAGTACAAGGCCACCGGGGGGCA 1845
QY 1741 GCCGTGTTCCAAAGCCGCTCAAGTTCCAGGTTGATATCACCTACACGAGGCTGGGGAG 1800
Db 1846 GCCGTGTTCCAAAGCCGCTCAAGTTCCAGGTTGATATCACCTACACGAGGCTGGGGAG 1905
QY 1801 GCGCAGAGGAGAACCGCATCTACTCCGTCACTTCCCTTCACTGCTCTCAGGCCCCAGCCCT 1860
Db 1906 GCGCAGAGGAGAACCGCATCTACTCCGTCACTTCCCTTCACTGCTCTCAGGCCCCAGCCCT 1965
QY 1861 CCGTTCAAGAGGCTGGAGACCATCCAGGCCAGCTGCTGAGCACACACACCGCCCT 1920
Db 1966 CCGTTCAAGAGGCTGGAGACCATCCAGGCCAGCTGCTGAGCACACACACCGCCCT 2025
QY 1921 GCGGCCACGACTTGTTCAGACACCACTACTCTATGGAATGATGACGGGGGCTTTCC 1980
Db 2026 GCGGCCACGACTTGTTCAGACACCACTACTCTATGGAATGATGACGGGGGCTTTCC 2085
QY 1981 AAATGTGGAATATCCCGAAAGTTAA 2007
Db 2086 AAATGTGGAATATCCCGAAAGTTAA 2112

RESULT 3
ABX13584
ID ABX13584 standard; cDNA; 3364 BP.
XX AC ABX13584;
XX DT 10-FEB-2003 (first entry)
XX DE Human cDNA encoding serine protein kinase KSE336-2.
XX KW Human; ss; gene; chromosome 11p15.5-pter; astrocytoma; meningioma;
KW pancreatic adenocarcinoma; insulin-dependent diabetes mellitus 2;
KW helicoid peripapillary choriorretinal degeneration; brain; pancreas;
KW Beckwith-Wiedemann syndrome; congenital hyperinsulinism; KSE336-2.
XX OS Homo sapiens.
XX PH Key
XX CDS Location/Qualifiers
FT 482..2239
FT /*tag= a
FT /product= "KSE336-2"
XX PN US6455292-B1.
XX PD 24-SEP-2002.
XX PF 16-AUG-2001; 2001US-0930181.
XX PR 16-AUG-2001; 2001US-0930181.
XX PA (ORIG-) ORIGENE TECHNOLOGIES INC.
XX SHU Y, Fan W, Kovacs XF, Zidanic M, Jay G;
XX WPI; 2003-066233/06.
XX P-PSDB; ABG72383.
XX New isolated polynucleotide coding without interruption for a human
PT KSE336 polypeptide useful for preventing or treating
PT diseases/conditions relating to brain and pancreas, e.g. meningioma,
PT insulin-dependent diabetes mellitus 2
XX Claim 4; Column 43-48; 34pp; English.
XX The invention relates to an isolated polynucleotide (its complement
CC or a sequence 99% similar to it) coding without interruption for a human
CC KSE336 polypeptide, a serine protein kinase, comprising the KSE336-1 and
CC KSE336-2 splice variants appearing as ABG72382 and ABG72383. Also
CC included is a method of identifying an agent that modulates the

expression of KSE336 in brain, pancreas, brain progenitor or pancreas progenitor cells comprising: (a) contacting a cell population comprising the cells with a test agent under conditions effective for the test agent to modulate the expression of KSE336; and (b) determining if the test agent modulates the expression of KSE336. The polynucleotides are useful as molecular targets or drug targets, and for detecting, diagnosing, staging, monitoring, prognosticating, preventing or treating diseases or conditions relating to brain and pancreas, such as astrocytoma, meningioma, pancreatic adenocarcinoma, insulin-dependent diabetes mellitus 2, helicobacter pylori, choroid plexus papilloma, Beckwith-Wiedemann syndrome or congenital hyperinsulinism. The method and polynucleotides are useful in research, diagnosis, drug discovery, therapy, clinical medicine, forensic science and pathology. The gene for KSE336 is located on chromosome 1p13.5-pter. The present sequence encodes the KSE366 variant KSE366-2.

Query Match	99.3%;	Score 1993;	DB 25;	Length 3364;
Best Local Similarity	99.8%;	Pred. No. 0;		
Matches 2007;	Conservative	0;	Mismatches	0; Indels 4; Gaps 1;
QY	1	ATGACATCGACGGGGAAGGACGGCGCGCGCAGCAGCGCAGTATGTTGGCCCTTACCGG	60	
DB	229	ATGACATCGACGGGGAAGGACGGCGCGCGCAGCAGCGCAGTATGTTGGCCCTTACCGG	288	
QY	61	CTGAGAGAAGACGCTGTGGGCAAGGGGCGACAGAGTCTGTGTGAAGCTTGGGGGTTCACTTGCCTC	120	
DB	289	CTGAGAGAAGACGCTGTGGGCAAGGGGCGACAGAGTCTGTGTGAAGCTTGGGGGTTCACTTGCCTC	348	
QY	121	ACCTGCCAGAAGGTGGCCATCAAGATCGTCAACCGTGAGAAGCTCAGCGAGTCGGTGCTG	180	
DB	349	ACCTGCCAGAAGGTGGCCATCAAGATCGTCAACCGTGAGAAGCTCAGCGAGTCGGTGCTG	408	
QY	181	ATGAAGGTGAGCGGGAGATCGCCATCCTGAAAGCTCATTTGAGCACCCCCACGTCCTAAAG	240	
DB	409	ATGAAGGTGAGCGGGAGATCGCCATCCTGAAAGCTCATTTGAGCACCCCCACGTCCTAAAG	468	
QY	241	CTGCA CGA CGTTTATGAAAAA CAAAAAATATTTT - - - GTACCTGTGTGTAGAACACGTGTC	296	
DB	469	CTGCACGACGTTTATGAAAAA CAAAAAATATTTTGTAGGTACCTGTGTGTAGAACACGTGTC	528	
QY	297	AGTGTGTGAGCTCTTCGACACTCTGTTGTAAGAGGGGAGGCTGACGCTTAGAGGAGGCTCG	356	
DB	529	AGTGTGTGAGCTCTTCGACACTCTGTTGTAAGAGGGGAGGCTGACGCTTAGAGGAGGCTCG	588	
QY	357	GAAGTTCTTCCGGCAGATCATCTCTGGGCTGGACCTTCTGCCACAGGCCACTCCATATGCCA	416	
DB	589	GAAGTTCTTCCGGCAGATCATCTCTGGGCTGGA CTCTCTGCCACAGGCCACTCCATATGCCA	648	
QY	417	CAGGGATCTGAACCTGAAAAACCTCTCTGTGGA CGAGAAGAACAACTCCGATCTCGAGA	476	
DB	649	CAGGGATCTGAACCTGAAAAACCTCTCTGTGGA CGAGAAGAACAACTCCGATCTCGAGA	708	
QY	477	CTTTGGCATGGCGTCCCTGCAAGTTGGCGACAGCTGTTGAGAGCACCACTCTGTGGTCCCC	536	
DB	709	CTTTGGCATGGCGTCCCTGCAAGTTGGCGACAGCTGTTGAGAGCACCACTCTGTGGTCCCC	768	
QY	537	CCA CTACGCTGCCCGAGGTGATCCCGGGGGAGAA GTATGACGGCCGGAAGGGGACGT	596	
DB	769	CCA CTACGCTGCCCGAGGTGATCCCGGGGGAGAA GTATGACGGCCGGAAGGGGACGT	828	
QY	597	GTGGAGCTGCGGCGTCACTCTGTTGCGCTTCTGTTGGTGGGGGCTCTGCCCTTCGACGATGA	656	
DB	829	GTGGAGCTGCGGCGTCACTCTGTTGCGCTTCTGTTGGTGGGGGCTCTGCCCTTCGACGATGA	888	
QY	657	CAACTTTCGACAGCTGCTGGAGAAGGTGAAGCGGGGCGTGTTCACATGCGCGCACTTTAT	716	
DB	889	CAACTTTCGACAGCTGCTGGAGAAGGTGAAGCGGGGCGTGTTCACATGCGCGCACTTTAT	948	
QY	717	CCGCCCCGA CTGCCAGAGTCTGTCTACGGGGCATGATCGAGGTGACCGCCGACCGCCGCT	776	
DB	949	CCGCCCCGA CTGCCAGAGTCTGTCTACGGGGCATGATCGAGGTGACCGCCGACCGCCGCT	1008	

Matches 1942; Conservative 0; Mismatches 4; Indels 0; Gaps 0;									
Qy	1	ATGACATCGACGGGAAAGGACGGCGCGCAGACGCGCAGATATGTTGGGCCCTACCGG	60						
Ds	1	ATGACATCGACGGGAAAGGACGGCGCGCAGACGCGCAGATATGTTGGGCCCTACCGG	60						
Qy	61	CTGAGAAGACGCTGGGCAAGGGCAGACAGAGTCTGTGAAGCTGGGGGTTCATCGGTC	120						
Ds	61	CTGAGAAGACGCTGGGCAAGGGCAGACAGAGTCTGTGAAGCTGGGGGTTCATCGGTC	120						
Qy	121	ACCTGCCAGAAGTGGCCATCAAGATGTCACCGTGAGAGCTCAGCGAGTGGTGCTG	180						
Ds	121	ACCTGCCAGAAGTGGCCATCAAGATGTCACCGTGAGAGCTCAGCGAGTGGTGCTG	180						
Qy	181	ATGAAGTGGAGCGGGAGATCGCCATCTGAAGCTCATTTAGACACCCCGCCTCTAAAG	240						
Ds	181	ATGAAGTGGAGCGGGAGATCGCCATCTGAAGCTCATTTAGACACCCCGCCTCTAAAG	240						
Qy	241	CTGACGACGTTTATGAAAACAAAATAATTTGTACTGTGTGTAGAACACGTTGCAGGT	300						
Ds	241	CTGACGACGTTTATGAAAACAAAATAATTTGTACTGTGTGTAGAACACGTTGCAGGT	300						
Qy	301	GGTGAGCTCTTCGACTACCTGTTGAGAGGGAGGCTGACGCTAAGAGGCTCGGAAG	360						
Ds	301	GGTGAGCTCTTCGACTACCTGTTGAGAGGGAGGCTGACGCTAAGAGGCTCGGAAG	360						
Qy	361	TTCTTCGGCAGATCATCTCGCTGCTGACTTCTGCCACAGCCACTCCATATGCCACAGG	420						
Ds	361	TTCTTCGGCAGATCATCTCGCTGCTGACTTCTGCCACAGCCACTCCATATGCCACAGG	420						
Qy	421	GATCTGAACCTGTAACCTCTCTGTGGAAGAGAAACAAATCTCGGATCGCAGACTTT	480						
Ds	421	GATCTGAACCTGTAACCTCTCTGTGGAAGAGAAACAAATCTCGGATCGCAGACTTT	480						
Qy	481	GGCATGGGCTCCCTGCAAGTGGGACAGCTGTGGAGACCACTGTGGTCCCCCACC	540						
Ds	481	GGCATGGGCTCCCTGCAAGTGGGACAGCTGTGGAGACCACTGTGGTCCCCCACC	540						
Qy	541	TACGCTGCCCGAGGTCATCCGGGGGAGAGATATGACGCGCGGAAAGCGGAGCTGTGG	600						
Ds	541	TACGCTGCCCGAGGTCATCCGGGGGAGAGATATGACGCGCGGAAAGCGGAGCTGTGG	600						
Qy	601	AGCTGGCGGTCATCTGTTGCTGCTGGTGGGGGCTCTGCCCTTCGACGATGACAAAC	660						
Ds	601	AGCTGGCGGTCATCTGTTGCTGCTGGTGGGGGCTCTGCCCTTCGACGATGACAAAC	660						
Qy	661	TTGGACAGCTGCTGGAGAAGTGAAGCGGGCGTGTTCACATGCCGCACTTTATCCCG	720						
Ds	661	TTGGACAGCTGCTGGAGAAGTGAAGCGGGCGTGTTCACATGCCGCACTTTATCCCG	720						
Qy	721	CCCGACTGCCAGAGTCTGCTACGGGGCATGATCGAGGTGGACGCGCAGCGGCTCACG	780						
Ds	721	CCCGACTGCCAGAGTCTGCTACGGGGCATGATCGAGGTGGACGCGCAGCGGCTCACG	780						
Qy	781	CTAGAGCATTCAGAAAACATATGTTATATAGGGGCAAGATGAGCCCGAACCCAGAG	840						
Ds	781	CTAGAGCATTCAGAAAACATATGTTATATAGGGGCAAGATGAGCCCGAACCCAGAG	840						
Qy	841	CAGCCCATTCCTCGCAAGTGCAGATCCGCTGCTGCCAGCTTGGAGGACATCGACCCC	900						
Ds	841	CAGCCCATTCCTCGCAAGTGCAGATCCGCTGCTGCCAGCTTGGAGGACATCGACCCC	900						
Qy	901	GACGTGCTGGACAGATGACCTCACTGGGCTGCTTTCCGAGACCGCAACAAGCTGTGAG	960						
Ds	901	GACGTGCTGGACAGATGACCTCACTGGGCTGCTTTCCGAGACCGCAACAAGCTGTGAG	960						
Qy	961	GACCTGCTGTCGAGAGGAGAAACAGAGAGATGATTTACTTCTCTCTGAGACCGG	1020						
Ds	961	GACCTGCTGTCGAGAGGAGAAACAGAGAGATGATTTACTTCTCTCTGAGACCGG	1020						
Qy	1021	AAAGAAAGGTACCCGAGCAGAGATGAGGACCTGCCCGCGGAAACGAGATAGACCT	1080						
Ds	1021	AAAGAAAGGTACCCGAGCAGAGATGAGGACCTGCCCGCGGAAACGAGATAGACCT	1080						

Qy	1081	CCCCGAAAGCTGTGTGACTCCCGGATGTCTGAACCGGCAACGCGCGCGGCGGCAACGC	1140						
Ds	1081	CCCCGAAAGCTGTGTGACTCCCGGATGTCTGAACCGGCAACGCGCGCGGCGGCAACGC	1140						
Qy	1141	AAATCATCGAGGTGTCTAGCGTGAAGCGGCGGCTCCCGGTGCTCGCGGCGGGCC	1200						
Ds	1141	AAATCATCGAGGTGTCTAGCGTGAAGCGGCGGCTCCCGGTGCTCGCGGCGGGCC	1200						
Qy	1201	ATTGAGATGCCCAGCAGCGCCAGAGTCTCGGTCTCATCAGCGGTGCTCTCAGGCTTT	1260						
Ds	1201	ATTGAGATGCCCAGCAGCGCCAGAGTCTCGGTCTCATCAGCGGTGCTCTCAGGCTTT	1260						
Qy	1261	TCCACAGCCCCACTCAGCAGCCCCCGGGTGAACCTCTCACCCCTCAACAAAGGGGAGTCC	1320						
Ds	1261	TCCACAGCCCCACTCAGCAGCCCCCGGGTGAACCTCTCACCCCTCAACAAAGGGGAGTCC	1320						
Qy	1321	CTCCCCACCCCAAGGGGACACTGTCTCCACCGCCAAAGAGAGCCCGGCTGGCAGCCC	1380						
Ds	1321	CTCCCCACCCCAAGGGGACACTGTCTCCACCGCCAAAGAGAGCCCGGCTGGCAGCCC	1380						
Qy	1381	AACCCCAACCCCGCTCCAGCCCGAGCGTGGAGGGGTGCTTGGAGGGCGCGGCTCAAC	1440						
Ds	1381	AACCCCAACCCCGCTCCAGCCCGAGCGTGGAGGGGTGCTTGGAGGGCGCGGCTCAAC	1440						
Qy	1441	TCCATCAAGAACAGCTTTCTGGGCTCACCCCGTTCACCGCGGAAACTGCAAGTTCCG	1500						
Ds	1441	TCCATCAAGAACAGCTTTCTGGGCTCACCCCGTTCACCGCGGAAACTGCAAGTTCCG	1500						
Qy	1501	ACGCCGAGGAGATGTCACCTGACACAGAGTGTCTCCCGAGAGCTGGGAGAGTCC	1560						
Ds	1501	ACGCCGAGGAGATGTCACCTGACACAGAGTGTCTCCCGAGAGCTGGGAGAGTCC	1560						
Qy	1561	TGGTTTGGGAACTTCATCAGCCCTGGAGAGGAGGAGCAGATCTTCGTGTCTCATCAAGAC	1620						
Ds	1561	TGGTTTGGGAACTTCATCAGCCCTGGAGAGGAGGAGCAGATCTTCGTGTCTCATCAAGAC	1620						
Qy	1621	AAACCTCTGAGCTCCATCAAGGCTGATCGTGCACGCTTCCTGTCTCATTCAGTCTC	1680						
Ds	1621	AAACCTCTGAGCTCCATCAAGGCTGATCGTGCACGCTTCCTGTCTCATTCAGTCTC	1680						
Qy	1681	AGCCAGAGCTCATCTCCCAACGAGCTTCGGGCGGAGTACAAAGGCCACGGGGGGCCA	1740						
Ds	1681	AGCCAGAGCTCATCTCCCAACGAGCTTCGGGCGGAGTACAAAGGCCACGGGGGGCCA	1740						
Qy	1741	GCGGTGTCCAGAAAGCGGTCAAAGTTCCAGGTTGATATCACTACAGGAGGTGGGGAG	1800						
Ds	1741	GCGGTGTCCAGAAAGCGGTCAAAGTTCCAGGTTGATATCACTACAGGAGGTGGGGAG	1800						
Qy	1801	GCGCAGAGGAGAAAGCGCATCTACTCGGTCACTTCACCTGTCTCAGGCCCCAGCGT	1860						
Ds	1801	GCGCAGAGGAGAAAGCGCATCTACTCGGTCACTTCACCTGTCTCAGGCCCCAGCGT	1860						
Qy	1861	CGCTTCAAGAGGTGGTGGAGACCTCCAGGCCCCAGTGTCTGAGCACAACGCCCCCT	1920						
Ds	1861	CGCTTCAAGAGGTGGTGGAGACCTCCAGGCCCCAGTGTCTGAGCACAACGCCCCCT	1920						
Qy	1921	GCGGCCAGCAGCTTGTTCAGACACAC	1946						
Ds	1921	GCGGCCAGCAGCTTGTTCAGACACAC	1946						

RESULT 6

ID	ABA02994
XX	ABA02994 standard; cDNA; 2217 BP.
AC	ABA02994;
XX	19-FEB-2002 (first entry)
XX	Human protein kinase 2246 encoding cDNA SEQ ID NO 1.
DE	Human; protein kinase 2246; cytosolic; immunomodulator; carcinoma;
XX	
KW	

anti-inflammatory; analgesic; cardiovascular; cancer; sarcoma;
cellular proliferation disorder; cellular differentiation disorder;
metastatic; hematopoietic disorder; leukaemia; immune disorder;
inflammatory disorder; arthritis; autoimmune disease; diabetes mellitus;
psoriasis; Crohn's disease; cardiovascular disease; virus; pain;
gene therapy; ss.

Homo sapiens.

Key Location/Qualifiers
CDS 51..2075

/*tag= a
/product= "protein kinase 2246"

WO200181588-A2.

01-NOV-2001.

25-APR-2001; 2001WO-US13784.

25-APR-2000; 2000US-199391P.

(MILL-) MILLENNIUM PHARM INC.

PI Meyers R;

WPI; 2002-049281/06.

P-PSDB; AAM47830.

New protein kinase nucleic acid and polypeptide molecules, designated 2246, useful for diagnosing, preventing or treating cancer or a cellular proliferation/differentiation disorders, e.g. carcinoma, sarcoma or leukaemias

Claim 1; Fig 1; 11pp; English.

The invention relates to the human protein kinase 2246 gene and the isolated encoded polypeptide with cytostatic, immunomodulator, anti-inflammatory, analgesic and cardiovascular activity. The 2246 nucleic acid and polypeptide are useful for diagnosing, preventing or treating a subject having cancer or a cellular proliferation and/or differentiation disorder or at risk of developing cancer or a cellular proliferation and/or differentiation disorder. In particular, the disorder includes carcinoma, sarcoma, metastatic or haematopoietic disorders (e.g. leukaemias) or cancers of the lung, breast, thyroid, head neck, prostate or genito-urinary tract. The 2246 nucleic acid and polypeptide are also useful for treating immune disorders, e.g. inflammatory (e.g. respiratory inflammation or arthritis), autoimmune disease (e.g. diabetes mellitus, psoriasis, Wegener's granulomatosis, Crohn's disease or Grave's disease), for treating cardiovascular diseases, endothelial cell disorder, viral diseases or pain. The nucleic acid and polypeptide are also useful for evaluating the efficacy of a treatment of cancer or a cellular proliferation and/or differentiation disorder. The nucleic acid is useful for gene therapy. The present sequence is that of the 2246 encoding cDNA.

Sequence 2217 BP; 476 A; 715 C; 682 G; 344 T; 0 other;

Query Match 96.6%; Score 1939.6; DB 24; Length 2217;

Best Local Similarity 99.8%; Pred. No. 0;

Matches 1942; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

1 ATGACATCGACGGGAAGACGCGCGCGCAGACGCGCAGTATGTTGGGCCCTTACCGG 60

51 ATGACATCGACGGGAAGACGCGCGCGCAGACGCGCAGTATGTTGGGCCCTTACCGG 110

61 CTGAGAGAACGCTGGCAAGGGGAGACAGGCTCTGTGAAGCTGGGGGTTCACTGGGTC 120

111 CTGAGAGAACGCTGGCAAGGGGAGACAGGCTCTGTGAAGCTGGGGGTTCACTGGGTC 170

121 ACCTGCCAGAGGTGGCCATCAAGATCGTCAACCGTGAAGAGCTCAGCGATCGGTGCTG 180

171 ACCTGCCAGAGGTGGCCATCAAGATCGTCAACCGTGAAGAGCTCAGCGATCGGTGCTG 230

Qy 181 ATGAAGGTGGAGCGGAGATCGCGATCTCTGAAGCTCATTTGAGCACCCCGCAGCTCTAAAG 240
Db 231 ATGAAGGTGGAGCGGAGATCGCGATCTCTGAAGCTCATTTGAGCACCCCGCAGCTCTAAAG 290
Qy 241 CTGCAGACGTTTTATGAAAACAAAAAATATTTGTACCTGTGTGTAGAACACGTTGTCAGGT 300
Db 291 CTGCAGACGTTTTATGAAAACAAAAAATATTTGTACCTGTGTGTAGAACACGTTGTCAGGT 350
Qy 301 GGTGAGCTCTTCGACTACCTGGTGAAGAGGGAGGCTGACGCTTAAGAGGCTTCGGAAG 360
Db 351 GGTGAGCTCTTCGACTACCTGGTGAAGAGGGAGGCTGACGCTTAAGAGGCTTCGGAAG 410
Qy 361 TTCTTCCGGCAGATCATCTCTGCGCTGGACTTTCTGCCACAGCCACTCCATATGCCACAGG 420
Db 411 TTCTTCCGGCAGATCATCTCTGCGCTGGACTTTCTGCCACAGCCACTCCATATGCCACAGG 470
Qy 421 GATCTGAACCTGAAACCTCTCTGTGGACGAGAACCAACATCCGCATCGCAGACTTTT 480
Db 471 GATCTGAACCTGAAACCTCTCTGTGGACGAGAACCAACATCCGCATCGCAGACTTTT 530
Qy 481 GGCATGGCGTCCCTGACAGGTTGGGACAGCCTGTTGGAGACCAGCTGTGGGTCCCCCAG 540
Db 531 GGCATGGCGTCCCTGACAGGTTGGGACAGCCTGTTGGAGACCAGCTGTGGGTCCCCCAG 590
Qy 541 TAGCCCTGCCCGCAGAGGTGATCCCGGGGAGAGTATGACCGCCGGAAGCGGAGCTGTGG 600
Db 591 TAGCCCTGCCCGCAGAGGTGATCCCGGGGAGAGTATGACCGCCGGAAGCGGAGCTGTGG 650
Qy 601 AGCTGCGCGCTCATCTCTGTTTCGCTTCTGTTGGGGCTCTGCCCTTCGACGATGACAAC 660
Db 651 AGCTGCGCGCTCATCTCTGTTTCGCTTCTGTTGGGGCTCTGCCCTTCGACGATGACAAC 710
Qy 661 TTGGCAGAGCTGCTGGAGAGGTGAAGCGGGGCGTGTTCACATGCCGCACCTTTATCCCG 720
Db 711 TTGGCAGAGCTGCTGGAGAGGTGAAGCGGGGCGTGTTCACATGCCGCACCTTTATCCCG 770
Qy 721 CCCGACTGCCAGAGTCTGCTACCGGGGATGATCGAGGTGACGCGCGCAGCGCGGCTCACG 780
Db 771 CCCGACTGCCAGAGTCTGCTACCGGGGATGATCGAGGTGACGCGCGCAGCGCGGCTCACG 830
Qy 781 CTAGAGCACATTCAGAACACATATGGTATATAGGGGGCAAGAAATAGCCCGAACACAGAG 840
Db 831 CTAGAGCACATTCAGAACACATATGGTATATAGGGGGCAAGAAATAGCCCGAACACAGAG 890
Qy 841 CAGGCCATTCTCGCAAGGTGCAGATCCGCTCGCTGCCAGCCTGGAGGACATCGACCCC 900
Db 891 CAGGCCATTCTCGCAAGGTGCAGATCCGCTCGCTGCCAGCCTGGAGGACATCGACCCC 950
Qy 901 GAGTGTCTGCACAGCATGCACTCACTGGGCTGCTTCCGAGACCGCAACAAAGCTGCTCAG 960
Db 951 GAGTGTCTGCACAGCATGCACTCACTGGGCTGCTTCCGAGACCGCAACAAAGCTGCTCAG 1010
Qy 961 GACTGTCTGTCGAGGAGGAGAACCCAGGAGAGATGATTTACTTCTCTCTCTGGACCGG 1020
Db 1011 GACTGTCTGTCGAGGAGGAGAACCCAGGAGAGATGATTTACTTCTCTCTCTGGACCGG 1070
Qy 1021 AAAGAAAGGTACCCGACCGCAGGAGGATGAGACTGCCCCCCCCCGGACGAGATAGACCT 1080
Db 1071 AAAGAAAGGTACCCGACCGCAGGAGGATGAGACTGCCCCCCCCCGGACGAGATAGACCT 1130
Qy 1081 CCCCGGAAGCGTGTGACTCCCGGATGCTGAACCGGACCGCAAGCGGGCGGCGGACGAGACGC 1140
Db 1131 CCCCGGAAGCGTGTGACTCCCGGATGCTGAACCGGACCGCAAGCGGGCGGCGGACGAGACGC 1190
Qy 1141 AAATCCATGGAGGTGCTCAGCTGACGAGCGCGGCTTCCCGGTGCTGTCGCGCGGCGGCC 1200
Db 1191 AAATCCATGGAGGTGCTCAGCTGACGAGCGCGGCTTCCCGGTGCTGTCGCGCGGCGGCC 1250
Qy 1201 ATTGAGATGSCCCAGCAGCGCCAGAGGTCTCGGTTCATCAGCGGTGCTCTCTCAGGCTTT 1260
Db 1251 ATTGAGATGSCCCAGCAGCGCCAGAGGTCTCGGTTCATCAGCGGTGCTCTCTCAGGCTTT 1310

1261 TCACACAGCCCACTCAGACGCCCGCGGTGACCCCTCACCCCTCACCAAGGGGCGAGTCCC 1320
1311 TCACACAGCCCACTCAGACGCCCGCGGTGACCCCTCACCCCTCACCAAGGGGCGAGTCCC 1370
1321 CTCCTCCACCCCAAGGGGACACTGTCTCCACACCCCAAGAGAGCCCGGTGCGACGCC 1380
1371 CTCCTCCACCCCAAGGGGACACTGTCTCCACACCCCAAGAGAGCCCGGTGCGACGCC 1430
1381 AACCCACAGCCCGCGGTGACCCCGGTGACCCCGGTGACCCCGGTGACCCCGGTGAC 1440
1431 AACCCACAGCCCGCGGTGACCCCGGTGACCCCGGTGACCCCGGTGACCCCGGTGAC 1490
1441 TCCTCAAGAACAGCTTCTGCGGTGACCCCGGTGACCCCGGTGACCCCGGTGACCC 1500
1491 TCCTCAAGAACAGCTTCTGCGGTGACCCCGGTGACCCCGGTGACCCCGGTGACCC 1550
1501 ACCTCGGAGAGATGTCACACCTGACACAGAGTGTGCTCCAGAGCTGCGGGAAGATCC 1560
1551 ACCTCGGAGAGATGTCACACCTGACACAGAGTGTGCTCCAGAGCTGCGGGAAGATCC 1610
1561 TGGTTGGGAACCTTCATCAGCTGGAGAGGAGGAGAGTCTTCTGTTGTCATCAAGAC 1620
1611 TGGTTGGGAACCTTCATCAGCTGGAGAGGAGGAGAGTCTTCTGTTGTCATCAAGAC 1670
1621 AAACCTCTGAGCTCCATCAAGGCTGACATCGTGCACGCTTCTGTTGTCATCAAGAC 1680
1671 AAACCTCTGAGCTCCATCAAGGCTGACATCGTGCACGCTTCTGTTGTCATCAAGAC 1730
1681 AGCCACAGCTGATCTCTCCAAAGAGTCTTCCGCGCGAGTACAAAGGCGCGGCGCA 1740
1731 AGCCACAGCTGATCTCTCCAAAGAGTCTTCCGCGCGAGTACAAAGGCGCGGCGCA 1790
1741 GCGGTGTTCCAGAGCGGTCAAGTTCAGTTCATACCTACACGAGGTTGGGGAG 1800
1791 GCGGTGTTCCAGAGCGGTCAAGTTCAGTTCATACCTACACGAGGTTGGGGAG 1850
1801 GCGCAGAGGAGACGCGATCTACTCGGTGACCTTACCTGCTTCCAGCCCGGCGCGT 1860
1851 GCGCAGAGGAGACGCGATCTACTCGGTGACCTTACCTGCTTCCAGCCCGGCGCGT 1910
1861 GCGTTCAGAGGTTGGTGGAGACCATCAGCCCGAGTGTGAGCAGACAGCCCGCT 1920
1911 GCGTTCAGAGGTTGGTGGAGACCATCAGCCCGAGTGTGAGCAGACAGCCCGCT 1970
1921 GCGGCCAGCAGCTTGTGAGACACAC 1946
1971 GCGGCCAGCAGCTTGTGAGAACCCCC 1996

RESULT 7
ID ABQ77626 standard; cDNA; 1827 BP.
XX AC ABQ77626;
XX
XX 21-OCT-2002 (first entry)
XX Human kinase NHP #2-encoding cDNA, SEQ ID NO:3.
XX Human; kinase; NHP; chromosome 11; signal transduction; cytoplasmic;
XX nuclear; brain; pituitary; hypothalamus; adipose; cerebellum;
XX adrenal gland; foetal lung; foetal brain; cancer; drug screening;
XX mental disorder; diagnostic reagent; clinical trial monitoring; cosmetic;
XX nutraceutical; mutation detection; gene expression analysis;
XX chromosome mapping; transgenic animal; cytostatic; gene therapy;
XX gene; ss.
XX Homo sapiens.
XX OS
XX Key Location/Qualifiers
XX CDS 1..1827
XX /*tag= a
XX /product= "Human kinase NHP #2"

XX WO200259287-A2.
XX 01-AUG-2002.
XX 22-JAN-2002; 2002WO-US01818.
XX 23-JAN-2001; 2001US-263378P.
XX (LEXI-) LEXICON GENETICS INC.
XX Turner CA, Mathur B;
XX WPI; 2002-599780/64.
XX P-PSDB; ABB09558.
XX Novel polynucleotide encoding human proteins sharing sequence
XX similarity with animal kinases, useful for drug screening, diagnosis,
XX in gene therapy of disorders and diseases e.g. cancer
XX
XX Claim 3; Page 38-39; 40pp; English.
XX The invention relates to novel human kinases (ABB09557, AAB09558),
XX referred to as NHPs, and to nucleic acids encoding them (ABQ77625,
XX ABQ77626). The NHPs have structural similarity with animal kinases
XX including serine-threonine kinases, carbon catabolite repressing kinases,
XX calcium/calmodulin-dependent protein kinases, and CAMP-dependent protein
XX kinases. The NHPs do not contain consensus signal sequences, indicating
XX that they may be cytoplasmic or nuclear proteins, and are thought to be
XX involved in signal transduction. Polynucleotides encoding NHPs were
XX obtained using human brain, pituitary, hypothalamus, adipose, cerebellum,
XX adrenal gland, foetal lung, and foetal brain cDNAs and primers derived
XX from human genomic DNA. The gene encoding the NHPs is located on
XX chromosome 11. NHP nucleotides and proteins are useful for treating
XX mental, biological or medical disorders including cancer, and for
XX screening compounds useful in the treatment of such conditions. They can
XX also be used as diagnostic reagents, in clinical trial monitoring and in
XX cosmetic and nutritional applications. NHP nucleotides can additionally
XX be used in the detection of disease-associated mutations, in the analysis
XX of gene expression, for mapping chromosome 11, for the recombinant
XX expression of NHPs, to generate transgenic animals, in gene therapy,
XX and as part of ribozyme and/or triple helix sequences useful in the
XX modulation of NHP gene expression. The present sequence represents cDNA
XX encoding a 608 amino acid NHP (#2).
XX
XX Sequence 1827 BP; 406 A; 581 C; 535 G; 305 T; 0 other;
XX
XX Query Match 91.0%; Score 1827; DB 24; Length 1827;
XX Best Local Similarity 100.0%; Pred. No. 0;
XX Matches 1827; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 181 ATGAAGGTGGAGCGGAGATCGGATCCTGAGAGTCAATGAGCACCACCGTCTTAAG 240
Db 1 ATGAAGGTGGAGCGGAGATCGGATCCTGAGAGTCAATGAGCACCACCGTCTTAAG 60
QY 241 CTGCACGAGCTTTATGAAAACAAAAATATTTGTACCTGGTGTAGAACACGTTGAGGT 300
Db 61 CTGCACGAGCTTTATGAAAACAAAAATATTTGTACCTGGTGTAGAACACGTTGAGGT 120
QY 301 GGTGAGCTTTGAGTACCTGTTGAGAGAGGGAGGCTGACGCTTAAGGGCTCGGAG 360
Db 121 GGTGAGCTTTGAGTACCTGTTGAGAGAGGGAGGCTGACGCTTAAGGGCTCGGAG 180
QY 361 TTCTTCCGGCAGATCATCTCTGCGCTGGACTTCTGCGCAGAGCCACTTCCATATGCCACAGG 420
Db 181 TTCTTCCGGCAGATCATCTCTGCGCTGGACTTCTGCGCAGAGCCACTTCCATATGCCACAGG 240
QY 421 GATCTGAAACCTGAAACCTCTCTGCTGGACGAGAGAAACAATCCGATTCGAGACTTT 480
Db 241 GATCTGAAACCTGAAACCTCTCTGCTGGACGAGAGAAACAATCCGATTCGAGACTTT 300
QY 481 GGCATGGCTCCCTGAGGTGGCGACAGCTTGGAGACCGAGCTGTGGTCCCCCACC 540

Db 301 GGCAATGCGCTCCCTGACAGTTGGCGACAGCCTGTTGGAGACAGCTGTGGGTCCCCCAC 360
Qy 541 TACGCTCCCGGAGGTGATCCGGGGGAGAAATGATGACGCGCGGAGAGCGGAGCTGTGG 600
Db 361 TACGCTCCCGGAGGTGATCCGGGGGAGAAATGATGACGCGCGGAGAGCGGAGCTGTGG 420
Qy 601 AGCTGCGGCGTCACTCTGTTCCGCTTGTGTGGGGGCTCTGCGCTTGGACGATGACAAC 660
Db 421 AGCTGCGGCGTCACTCTGTTCCGCTTGTGTGGGGGCTCTGCGCTTGGACGATGACAAC 480
Qy 661 TTGCGACAGCTGCTGGAGAAAGTGAAGCGGGCGTGTTCACATGCGCGCACTTTATCCCG 720
Db 481 TTGCGACAGCTGCTGGAGAAAGTGAAGCGGGCGTGTTCACATGCGCGCACTTTATCCCG 540
Qy 721 CCCGACTGCCAGAGTCTCTACGGGGCATGATCAGAGTGAAGCGCGCATGCGGCTCAAG 780
Db 541 CCCGACTGCCAGAGTCTCTACGGGGCATGATCAGAGTGAAGCGCGCATGCGGCTCAAG 600
Qy 781 CTAGAGCAATTCAGAAACACATATGGTATATAGGGGCAAGATGAGCCCGAACCAGAG 840
Db 601 CTAGAGCAATTCAGAAACACATATGGTATATAGGGGCAAGATGAGCCCGAACCAGAG 660
Qy 841 CAGCCATCTCCGCAAGGTGAGATCCGCTGCTGCGCGCGCATGAGGAGATGACGCC 900
Db 661 CAGCCATCTCCGCAAGGTGAGATCCGCTGCTGCGCGCGCATGAGGAGATGACGCC 720
Qy 901 GACGTGCTGGACAGCATGACCTCACTGGGCTGCTTCCGAGACCGCAACAGCTGCTGAG 960
Db 721 GACGTGCTGGACAGCATGACCTCACTGGGCTGCTTCCGAGACCGCAACAGCTGCTGAG 780
Qy 961 GACCTGCTGCTCGAGGAGAGAACACAGGAGATGATTTACTTCTCTCTCTGAGCCGG 1020
Db 781 GACCTGCTGCTCGAGGAGAGAACACAGGAGATGATTTACTTCTCTCTCTGAGCCGG 840
Qy 1021 AAGAAAGGTATCCCGAGCAGGAGATGAGACCTGCCCCCGGAAACGAGATGACCT 1080
Db 841 AAGAAAGGTATCCCGAGCAGGAGATGAGACCTGCCCCCGGAAACGAGATGACCT 900
Qy 1081 CCCCGAAGCGTGTGACTTCCCGGATGCTGAACCGGACCGCAAGCGCGGCCAGAACGC 1140
Db 901 CCCCGAAGCGTGTGACTTCCCGGATGCTGAACCGGACCGCAAGCGCGGCCAGAACGC 960
Qy 1141 AATCATGAGGAGTCTCAGGTGACGACGCGCGCTCCCGGTGCTGCGCGCGCGGCC 1200
Db 961 AATCATGAGGAGTCTCAGGTGACGACGCGCGCTCCCGGTGCTGCGCGCGCGGCC 1020
Qy 1201 ATTGAGATGCGCCAGCAGCGCGGAGGTCTCGGTTCATCAGGCTGCTCTCAGGCTT 1260
Db 1021 ATTGAGATGCGCCAGCAGCGCGGAGGTCTCGGTTCATCAGGCTGCTCTCAGGCTT 1080
Qy 1261 TCCACAGCCCACTCAGCAGCCCCCGGTGACCCCTCAACCCCTCAACAGGGGAGTCCC 1320
Db 1081 TCCACAGCCCACTCAGCAGCCCCCGGTGACCCCTCAACCCCTCAACAGGGGAGTCCC 1140
Qy 1321 CTCGCCACCCCAAGGGAGACCTGTTCACACGCGCAAGAGAGAGCCCGGTGCGACGCC 1380
Db 1141 CTCGCCACCCCAAGGGAGACCTGTTCACACGCGCAAGAGAGAGCCCGGTGCGACGCC 1200
Qy 1381 AACCCACGCCCCCTCAGCCCCAGCGTCCGAGGGGTGCGGTGAGGGCGCGGTCAAC 1440
Db 1201 AACCCACGCCCCCTCAGCCCCAGCGTCCGAGGGGTGCGGTGAGGGCGCGGTCAAC 1260
Qy 1441 TCCATCAAGAACAGCTTTCTGGGCTCACCCGCTTCCAGCCCGGAAACTGCAAGTTCCG 1500
Db 1261 TCCATCAAGAACAGCTTTCTGGGCTCACCCGCTTCCAGCCCGGAAACTGCAAGTTCCG 1320
Qy 1501 AGCGCGGAGGAGATTCACCTGACACAGAGTCTCTCCAGAGCTGCGGAGAGTCC 1560
Db 1321 AGCGCGGAGGAGATTCACCTGACACAGAGTCTCTCCAGAGCTGCGGAGAGTCC 1380
Qy 1561 TGGTTTGGAACTTTCATCAGCTTGGAGAGGAGGAGGAGTCTTCTGCTGATCAAGAC 1620
Db 1381 TGGTTTGGAACTTTCATCAGCTTGGAGAGGAGGAGGAGTCTTCTGCTGATCAAGAC 1440

Qy 1621 AAACCTCTGAGTCCCATCAAGGCTGACATCGTCACGCCCTTCTGTCGATTCCAGTCTC 1680
Db 1441 AAACCTCTGAGTCCCATCAAGGCTGACATCGTCACGCCCTTCTGTCGATTCCAGTCTC 1500
Qy 1681 AGCCACAGCGTCACTCTCCAAACGAGCTTCCGGGCCAGGTACAAGGCCACCGGGGGCCA 1740
Db 1501 AGCCACAGCGTCACTCTCCAAACGAGCTTCCGGGCCAGGTACAAGGCCACCGGGGGCCA 1560
Qy 1741 GCGTGTTCAGAGCGGCTCAAGTTCCAGGTTGATATCACTACCGAGGGTGGGGAG 1800
Db 1561 GCGTGTTCAGAGCGGCTCAAGTTCCAGGTTGATATCACTACCGAGGGTGGGGAG 1620
Qy 1801 GCGCAGAGGAGAACGCGCATCTACTCGTCACCTTCACTCCTGCTCAGGCCCCAGCCGT 1860
Db 1621 GCGCAGAGGAGAACGCGCATCTACTCGTCACCTTCACTCCTGCTCAGGCCCCAGCCGT 1680
Qy 1861 CGTTCAAGAGGCTGTGGAGACCATCCAGGCCAGCTGCTGAGCACACAGACCCGCCCT 1920
Db 1681 CGTTCAAGAGGCTGTGGAGACCATCCAGGCCAGCTGCTGAGCACACAGACCCGCCCT 1740
Qy 1921 GCGGCCAGCAGCTTTCAGACACCATTAAGTATGAAATGATGAGGGCGGCTTTCC 1980
Db 1741 GCGGCCAGCAGCTTTCAGACACCATTAAGTATGAAATGATGAGGGCGGCTTTCC 1800
Qy 1981 AATGTGGAATTATCCCGAAAGTTAA 2007
Db 1801 AATGTGGAATTATCCCGAAAGTTAA 1827

RESULT 8

AA151889
ID AA151889 standard; DNA; 2799 BP.
XX
AC AA151889;
XX
DT 08-MAY-2003 (first entry)
XX
DE Human cell cycle-regulatory factor Cdr2 coding sequence.
XX
KW Human; cell cycle-regulatory factor; Cdr2; kinase; proliferative disease;
XX anticancer agent; wound-healing drug; gene; ds.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT CDS 50..2314
FT /*tag= a
FT /product= "Human Cdr2 protein"
XX
PN W0200299110-A1.
XX
PD 12-DEC-2002.
XX
XX 03-JUN-2002; 2002WO-JP05411.
XX
XX 04-JUN-2001; 2001JP-0168792.
XX
PA (TAIH) TAIHO PHARM CO LTD.
XX (NAKA/) NAKANISHI M.
XX
XX Nakanishi M;
XX
XX WPI; 2003-156857/15.
XX P-PSDB; AAO16604.
XX
XX Cell cycle-regulatory factor Cdr2 with kinase activity and encoded
XX gene, applicable in diagnosis of and screening drugs for proliferative
XX diseases e.g. anticancer agents and wound-healing drugs
XX
XX Claim 2; Fig 2 A-C; 63pp; Japanese.
XX
XX The invention comprises the amino acid and coding sequence of the human

Db	2103	TCGAGGCACAGCTCCTGAGCACTCATGACACAGCCCTCCGTGTCAGGCCCTTGGCAGACGAGA	2163
Qy	1946	CTAACTG 1952	
Db	2163	AGAACGG 2169	
RESULT 9			
ABV74557	ID	ABV74557 standard; cDNA; 2337 BP.	
XX	AC	ABV74557;	
XX	XX	20-JAN-2003 (first entry)	
DT	XX	Human kinase #1 coding sequence.	
DE	XX	Human; kinase; chromosome 19; gene; ss.	
XX	XX	Homo sapiens.	
XX	XX	Key Location/Qualifiers	
FH	FT	1..2337	
CDS	FT	/*tag= a	
FT	FT	/product= "Human kinase #1"	
XX	XX	WO200281670-A1.	
PN	XX	17-OCT-2002.	
PD	XX	04-APR-2002; 2002WO-US10786.	
PF	XX	06-APR-2001; 2001US-282036P.	
PR	XX	(LEXI-) LEXICON GENETICS INC.	
PA	XX	Turner CA, Mathur B, Friddle CJ;	
PI	XX	WPI; 2003-058538/05.	
XX	XX	P-PSDB; ABB98743.	
XX	XX	New human kinase proteins useful for diagnosis, drug screening,	
PT	PT	clinical trial monitoring, treatment of disorders and diseases, and	
PT	PT	cosmetic and nutritional applications	
XX	XX	Claim 5; Page 39; 47pp; English.	
XX	XX	The present sequence is the coding sequence for a novel human kinase. The	
CC	CC	genomic locus encoding the kinase is thought to be on human chromosome	
CC	CC	19. The kinase and its coding sequence are useful for diagnosis, drug	
CC	CC	screening, clinical trial monitoring, treatment of disorders and	
CC	CC	diseases, and cosmetic and nutritional applications.	
XX	XX	Sequence 2337 BP; 479 A; 786 C; 686 G; 386 T; 0 other;	
SQ	XX	Query Match 48.1%; Score 965; DB 25; Length 2337;	
		Best Local Similarity 70.0%; Pred. No. 1.5e-199;	
		Matches 1484; Conservative 0; Mismatches 440; Indels 195; Gaps 6;	
Qy	29	CGCAGCAGCGCAGTATGTTGGGCCCTTACCGGCTGGAGAAAGACGCTGGCAAGGGGCGAGA 88	
Db	74	CCCAGCAGCGCCCAATATGTGGGCCCTTATCGGCTGGAGAAAGACGCTGGCAAGGACAGA 133	
Qy	89	CAGGTCGTGTGAAGCTGGGGGTTTCACTGCTCACCTGCCAGAGGTGGCCATCAAGATCG 148	
Db	134	CAGGTCGTGTAACTCGGGGTCCATGTGCATCAGGTCAGAAGGTGCCATCAAGATCG 193	
Qy	149	TCAACCGTGAAGAAGCTCAGCGAGTTCGGTGTGTATGAAGGTGGAGCGGGAGATCGCCATCC 208	
Db	194	TGAACCGGAGAAAGCTGTCGAGTTCGGTGTGTATGAAGGTGGAGCGGGAGATCGCCATCC 253	
Qy	209	TGAAGCTCATTTAGACACCCGCCCAAGCTTAAAGCTGCACGACGTTTATGAAAAACAAAAAT 268	

QY 1290 ----- 1289
Db 1394 CTGAGGCGGGGCTCCCGACTTCCAAAACGACAGCTGCTTCTCGGGGCCCCAGGG 1453
QY 1290 -----GACCCCTCACCCCTCAACGAGGCGAGTCCCTCCCTCC----- 1326
Db 1454 GTGGGGCGCGGGAGCAGCCCGCGCCCAAGTGCCTCCACACCCCTCCCGGCG 1513
QY 1327 -----ACCCCAAGGGGACACCTGTCCACACGCCAA 1357
Db 1514 CCCCAGCTCCCGCGCTCTCTGCGGGGACCCCTTGCACTCGCTCTGCAACGCCCC 1573
QY 1358 AGGAGAGCCCGGCTGGCAGCCCAACCCCGCCCGCTCCAGCCCC-----AGGCTCG 1411
Db 1574 GGCCAGTCCACCGGACCCCGGGGACACACACCCCGCGCGGTGGCTCG 1633
QY 1412 GAGGGTGCCTGAGGGCGGCTCAACTCCATCAAGAAACAGTTTCTGGGTCAACCC 1471
Db 1634 GGGAGCGGCTGGAGAGTCTCTCAACTCCATCCGCAACAGCTTCTGGGTCCCTC 1693
QY 1472 GCTTCCACCGCGGAACTGCAAGTTCGAGCGCGGAGGAGATGTCCAACTGACACCAG 1531
Db 1694 GCTTTCACCGCGCAAGATGAGGTCTCCCTACCGCTGAGGAGATGTCCAGCTTGACGCCAG 1753
QY 1532 AGTCGTCCCGAGAGCTGGCGAAGAGTCTCTGGTTTGGAACTTTCATCAGCGCTGGAGAAGG 1591
Db 1754 AGTCCTCCCGAGAGCTGGCAAAACGCTCTCTGGTTTGGGAACTTTCATCTCTTGACAAAG 1813
QY 1592 AGGAGAGATCTTCGTGGTCAACAAAGACAAACCTCTGAGTCTCATCAAGGCTGACATCG 1651
Db 1814 AAGAACAAATATCTCTGTGCTAAAGGACAAACCTCTCAGCAGCATCAAGAGCAGACATCG 1873
QY 1652 TGCAAGCTTCTCTCGATTCCTAGTCTTCAGCCACAGCGTATCTCCCAACGAGCTTCC 1711
Db 1874 TCCATGCTTCTCTCGATTCCTAGTCTTCAGTCAAGTCTGCTGACACACAGCTTCA 1933
QY 1712 GGGCGAGTCAAGGCCCAAGGGGGGCGCAGCGTGTTCAGAAAGCGGTCAAGTTCCAGG 1771
Db 1934 GGGCGAGTCAAGGCCCAAGTGGCGGCGCTCTCGTCTCCAAAGCCGCTCCGCTTCCAGG 1993
QY 1772 TTGATATCACTACACGAGGCT-----GGGAGGCGCAGAGAGAGA 1813
Db 1994 TGGACATCAGCTCTCTGAGGGTTCAGAGCCCTCCCGCGAGCGGACGCGAGGGTG 2053
QY 1814 ACGGATCTACTCGTCACTTCACTTCACTCTCAGCGCCCGAGCGTCTTCAAGAGGG 1873
Db 2054 GTGGCACTCTCTCGTCACTTCACTCTCTCGGTCCAGCGCTCGGTTCAGCGAG 2113
QY 1874 TGGTGGAGACCATCCAGGCCAGCTCTGAGCAGACACAGCCCGCTGCGGCCAGCACT 1933
Db 2114 TGGTGGAGACCATCCAGGCCAGCTCTGAGCAGACACTCATGACAGCCCTCGTGCAGGCC 2173
QY 1934 TGTACAGACCACTAACTG 1952
Db 2174 TGGCAGACGAGAAGACGG 2192

RESULT 10

AAS06717
ID AAS06717 standard; cDNA; 2385 BP.
XX AC AAS06717;
XX DT 12-SEP-2001 (first entry)
XX DE Polynucleotide sequence encoding human protein kinase #17.
XX KW Human; protein kinase; PTK; STK; cancer; cardiovascular disease;
KW metabolic disorder; immune related disease; neurological disorder;
KW neurodegenerative disorder; inflammatory disorder; infectious disease;
KW reproductive disorder; gene therapy; ss.
XX

OS Homo sapiens.
XX WO200138503-A2.
PN 31-MAY-2001.
XX 22-NOV-2000; 2000WO-US32085.
PF 24-NOV-1999; 99US-0167482.
XX (SUGEN-) SUGEN INC.
PA Plowman GD, Whyte D, Manning G, Sudarsanam S, Martinez R;
PI Flanagan P, Clary D;
XX WPI: 2001-343950/36.
DR P-PSDB; AAU03517.
XX Nucleic acids encoding human kinase polypeptides, useful for preventing
PT diagnosing and/or treating e.g. cancer, immune, cardiovascular and
PT neuronal-associated diseases, and microbial infections -
XX Example 1; Figure 1; 433pp; English.
PS AAS06701-AAS06757 encode for novel human protein kinases #1-57. The
XX novel protein kinases have been identified as members of the tyrosine
CC or serine/threonine kinase (PTK and STK) families. The polynucleotides
CC encoding protein kinases and the polypeptides may be used in the
CC prevention, diagnosis and treatment of diseases associated with
CC inappropriate kinase expression. For example, they may be used to treat
CC cancers (especially cancers of haematopoietic origin), cardiovascular
CC disease (e.g. atherosclerosis), metabolic disorders (e.g. diabetes),
CC immune related diseases (e.g. rheumatoid arthritis), neurological
CC disorders (e.g. schizophrenia), neurodegenerative disorders (e.g.
CC Parkinson's disease), inflammatory disorders (e.g. asthma), infectious
CC disease (e.g. HIV) and reproductive disorders (e.g. infertility).
CC Additionally, polynucleotides encoding protein kinases may be
CC used for gene therapy and as DNA probes in diagnostic assays.
CC The protein kinase polypeptides may be used as antigens in the production
CC of antibodies against the protein kinases and in assays to identify
CC modulators of protein kinase expression and activity.
XX SQ Sequence 2385 BP; 494 A; 768 C; 731 G; 392 T; 0 other;
Query Match 48.1%; Score 964.4; DB 22; Length 2385;
Best Local Similarity 69.8%; Pred. No. 2e-199;
Matches 1490; Conservative 0; Mismatches 451; Indels 195; Gaps 6;
QY 12 GGGGAAGGACGGCGCGCGAGCAGCGCAGTATGTTGGGCCCTACCGCTGGAGAAGAC 71
Db 105 GGAGGCCGAAGAGAGAGAGGGCGCCACGCCCAATATGTGGGCCCTATCGGCTGGAGAAGAC 164
QY 72 GCTGGCCAGAGGGGACACAGGCTGTGCTGAAGCTGGGGGTTCACTGCTCACCTGCCAGAA 131
Db 165 GCTGGCCAAAGGACACAGAGGCTGTTAACTCGGGGTTCCACTGATCATCGGCTCAGAA 224
QY 132 GGTGGCCATCAAGATCGTCAACCGTGAAGAGCTCAGCGAGTGGTGTCTGATGAAGTGA 191
Db 225 GGTGCCCATCAGATCGTGAACCGGAGAGGCTGTCCGAGTGGTGTCTGATGAAGTGA 284
QY 192 GCGGAGATCGCGATCTCTGAGCTATTGAGCAACCCCAAGTCTTAAGCTGACACAGT 251
Db 285 GCGGAGATCGCCATCTGAAGCTCATCGAACACCCACATGCTCTCAAGCTCCACAGCT 344
QY 252 TTATGAACCAAAAAATATTTGTACCTGGTGTGCTGAGAACAGCTGTGAGGTGGTGTCTT 311
Db 345 CTACGAGAACAGAAATATTTGTACCTGGTGTGAGACAGCTGTGAGGTGGTGTGATTT 404
QY 312 CGACTACCTGGTGAAGAGGGAGGCTGACGCTTAAGAGGCTCGGAAGTTCTTCCGGCA 371
Db 405 CGACTACCTGGTGAAGAGGGAGAGCTGACGCCCAAGAGGCCCGGAAAGTTCTTCCGCCA 464
QY 372 GATCATCTCTGCGCTGAGACTTTCTGCGACAGCCACTCATATATGCCACAGGGATCTGAAC 431

[illegible]

Accession	Species	Gene	Position	Sequence	Length
1545	Ca	ACACCCCTGCCCGGCCCC	1545	CAACACCCCTGCCCGGCCCC	1604
1341	Ac	CTGTCCACAGCCAAAGAGAGAGCCCGGCTGGCAGCGCCCAACCCACGCGCCCGCTCCAG	1341	ACCTGTCCACAGCCAAAGAGAGAGCCCGGCTGGCAGCGCCCAACCCACGCGCCCGCTCCAG	1400
1605	Gc	CTGTGACACAGCCCGCGGCGAGTCCACCGGGACCCCGGGGACAAACACCCCGCCAG	1605	GCTCTGTGACACAGCCCGCGGCGAGTCCACCGGGACCCCGGGGACAAACACCCCGCCAG	1664
1401	Cc	CCCC-----AGCGTCGGAGGGGTGCCCTGGAGGGCGCGGCTCAATCCATCAAGAACAG	1401	CCCC-----AGCGTCGGAGGGGTGCCCTGGAGGGCGCGGCTCAATCCATCAAGAACAG	1454
1665	Cc	CGCGCGGTGGCGTCCGGGGAGCGCCTCGGAGGAGTCTCAATCCATCCATCGCCACACAG	1665	CGCGCGGTGGCGTCCGGGGAGCGCCTCGGAGGAGTCTCAATCCATCCATCGCCACACAG	1724
1455	Ct	TTTCTGGGCTCACCCCGCTTCCACCGCGCGAAATCTCAAGTTCCGACCGCGGAGGAT	1455	TTTTTCTGGGCTCACCCCGCTTCCACCGCGCGAAATCTCAAGTTCCGACCGCGGAGGAT	1514
1725	Ct	TTTCTGGGCTCCCCCTCGCTTTTCCCGCGGCAAGATGCAAGTCCCTACCGCTGAGGAGAT	1725	TTTTTCTGGGCTCCCCCTCGCTTTTCCCGCGGCAAGATGCAAGTCCCTACCGCTGAGGAGAT	1784
1515	Gt	CAACCTGTACACACAGAGTCGTCCTCCAGAGCTGGCGAAGAACTCTCTGGTTTGGGAACCT	1515	GTCAACCTGTACACACAGAGTCGTCCTCCAGAGCTGGCGAAGAACTCTCTGGTTTGGGAACCT	1574
1785	Gt	CACGCTTTCAGCCAGAGTCTCTCCCGGAGCTGGCANAACGCTCTCTGGTTTCGGGAACCT	1785	GTCCAGCTTTCAGCCAGAGTCTCTCCCGGAGCTGGCANAACGCTCTCTGGTTTCGGGAACCT	1844
1575	Ct	CAGCCTGGAGAAAGGAGGAGCAGATCTTTCGTTGGTTCATAAAGACAAACCTCTGAGCTC	1575	CATCAGCCTGGAGAAAGGAGGAGCAGATCTTTCGTTGGTTCATAAAGACAAACCTCTGAGCTC	1634
1845	Ct	CTCTCTTTGGACAAAGAAAGAACAAATATTTCTCTGCTAAAGGACAAACCTCTCAGCAG	1845	CATCTCTTTGGACAAAGAAAGAACAAATATTTCTCTGCTAAAGGACAAACCTCTCAGCAG	1904
1635	Ct	CAGGCTGACATCGTGCAGCGCTTCTCTGTCTGATTCAGGCTCTCAGCCACAGCGTCAT	1635	CATCAGGCTGACATCGTGCAGCGCTTCTCTGTCTGATTCAGGCTCTCAGCCACAGCGTCAT	1694
1905	Ct	CAAAAGCAGACATCGTCCATGCCCTTCTGTCTGATCCCGCAGCTCAGTACAGTGTGCT	1905	CATCAAAAGCAGACATCGTCCATGCCCTTCTGTCTGATCCCGCAGCTCAGTACAGTGTGCT	1964
1695	Ct	CCCAAAACGAGCTTCCGGCCGAGTACAAGGCCACCGGGGGCCAGCCGCTTCCCGGACG	1695	CTCCCAAAACGAGCTTCCGGCCGAGTACAAGGCCACCGGGGGCCAGCCGCTTCCCGGACG	1754
1965	Gt	CACAGACGCTTTCAGGGCCGAGTACAAGGCCAGTGGCGGCCCTCTCGTCTTCANAA	1965	GTCAACAGACGCTTTCAGGGCCGAGTACAAGGCCAGTGGCGGCCCTCTCGTCTTCANAA	2024
1755	Gc	CGGTCAATTCCAGGTTGATATCACTTACCTACACGGAGGT-----GG	1755	GCCGCTCAATTCCAGGTTGATATCACTTACCTACACGGAGGT-----GG	1796
2025	Gc	CGGTCCGCTTCCAGTGGACATCATGCTCTCTGAGGGTCCAGAGCCCTCCCGGACG	2025	GCCCGTCCGCTTCCAGTGGACATCATGCTCTCTGAGGGTCCAGAGCCCTCCCGGACG	2084
1797	Gg	AGGCGCAGAGGAGAAAGGCACTTACTCCGTCATCTTCAACCTCTCTCAGGCCCCCAG	1797	GGAGGCGCAGAGGAGAAAGGCACTTACTCCGTCATCTTCAACCTCTCTCAGGCCCCCAG	1856
2085	Gg	AGGCGCAGGAGGTGGTGGCATCTACTCCGTCATCTTCACTCTCTCAGTCTCAGT	2085	GGAGGCGCAGGAGGTGGTGGCATCTACTCCGTCATCTTCACTCTCTCAGTCTCAGT	2144
1857	Cc	CGCTCTCAAGAGGGTGGTGGAGACCATCCAGGCCCCAGCTCTGAGCACACACGACCC	1857	CCGCTCGCTTCAAGAGGGTGGTGGAGACCATCCAGGCCCCAGCTCTGAGCACACACGACCC	1916
2145	Cc	CGTCTGTTCAAGGAGTGGTGGAGACCATCCAGGCGACAGCTCTGAGCACTATGACCA	2145	CCGTCTGTTCAAGGAGTGGTGGAGACCATCCAGGCGACAGCTCTGAGCACTATGACCA	2204
1917	Gc	TGCGGCCAGCCTTGTTCAGACACCACTAACTG	1917	GCTTGTGCGGCCAGCCTTGTTCAGACACCACTAACTG	1952
2205	Gc	CCGTCCGTGCAGGCCCTGGCAGACGAGAAGACGG	2205	GCCCTCCGTGCAGGCCCTGGCAGACGAGAAGACGG	2240
RESULT 11					
AAD26464					
ID	AAD26464	standard; cDNA; 2897 BP.			
XX					
AC	AAD26464;				
XX					
XX	26-MAR-2002	(first entry)			
XX					
XX	Human kinase PKIN-17	cdNA.			
KW	Human; kinase; PKIN-17; cancer; leukaemia; adenocarcinoma; osteoporosis;				
KW	immune disorder; atherosclerosis; Crohn's disease; Hodgkin's disease;				
KW	Acquired Immune Deficiency Syndrome; AIDS; Addison's disease; anaemia;				
KW	allergy; asthma; adult respiratory distress syndrome; multiple sclerosis;				
KW	autoimmune thyroiditis; bronchitis; diabetes mellitus; osteoarthritis;				
KW	Good pasture's syndrome; Graves' disease; pancreatitis; psoriasis;				
KW	rheumatoid arthritis; ulcerative colitis; cirrhosis; Cushing's syndrome;				
KW	hepatitis; hypothyroidism; cerebral palsy; cataract; angina pectoris;				
KW	cardiovascular disease; hypertension; vasculitis; myocarditis; obesity;				
KW	congestive heart failure; ischaemic heart disease; lung tumour; gout;				
XX	fatty liver; Niemann-Pick's disease; gene therapy; ss.				
OS	Homo sapiens.				
XX					
Key	Location/Qualifiers				
FT	1..2385				
CD					

RESULT 11

AAD26464

ID AAD26464 standard; cDNA; 2897 BP.

XX

AC AAD26464;

XX

26-MAR-2017

XX

DE Human kin

X

KW Human; ki

KW immune di

KW Acquired 11/1/2000

allergy;

KW autoimmune

KW	Good	past
KW	rheumat	of

KW pneumato

KW	nepācīti
KW	cardi ovag

KW
cardiovasc
KW
constitutiv

KW congestive
KW fatty liv

XX
racy 11v

/*tag= a
/product= "Human PKIN-17 protein"

FT WO200196547-A2.
 XX 20-DEC-2001.
 XX 14-JUN-2001; 2001WO-US19444.
 XX 15-JUN-2000; 2000US-212073P.
 XX 23-JUN-2000; 2000US-213467P.
 XX 30-JUN-2000; 2000US-215651P.
 XX 07-JUL-2000; 2000US-216605P.
 XX 13-JUL-2000; 2000US-218372P.
 XX 25-AUG-2000; 2000US-228056P.
 XX (INCYTE) INCYTE GENOMICS INC.
 XX Yue H, Lal P, Bandman O, Borowsky ML, Au-Young J, Lu Y;
 PI Gandhi AR, Tribouley CM, Walia NK, Yao MG, Lu DM, Greenwald SR;
 PI Ramkumar J, Griffin JA, Kearney L, Burford N, Nguyen DB, Tang VT;
 PI Baughn MR, He A, Thornton M, Hafalia A, Patterson C, Gururajan R;
 PI Lo TP, Khan F, Recipon SA, Azimzai Y, Policky JL, Ding L;
 PI Grether M, Elliott VS, Thangavelu K, Batra S, Ison CH;
 XX WPI; 2002-090207/12.
 DR P-PSDB; AAE16271.
 XX New polypeptides, useful for diagnosing, treating or preventing
 PT disorders of growth and development, cardiovascular and lipid, and
 PT diseases such as cancer, comprise human kinase polypeptides -
 XX Claim 5; Page 188-189; 197pp; English.
 XX The invention relates to human kinase PKIN proteins and their
 CC corresponding cDNAs. A composition containing PKIN agonist is useful for
 CC treating a disease or condition associated with decreased expression of
 CC PKIN and a composition comprising PKIN antagonist is useful for treating
 CC a disease or condition associated with overexpression of PKIN. The
 CC diseases include cancer (leukaemia, adenocarcinoma, lymphoma, melanoma,
 CC myeloma, sarcoma, teratocarcinoma, Hodgkin's disease); immune disorder
 CC (Acquired Immune Deficiency Syndrome (AIDS), asthma, Addison's disease,
 CC atherosclerosis, anaemia, allergies, adult respiratory distress syndrome,
 CC autoimmune thyroiditis, gout, bronchitis, Crohn's disease, diabetes
 CC mellitus, multiple sclerosis, Good pasture's syndrome, Graves' disease,
 CC osteoarthritis, osteoporosis, pancreatitis, psoriasis, Reiter's syndrome,
 CC rheumatoid arthritis, Sjogren's syndrome, uveitis, ulcerative colitis,
 CC bacterial, parasitic, fungal, viral, protozoal and helminthic infections)
 CC growth and development disorders (arteriosclerosis, cirrhosis, hepatitis,
 CC Cushing's syndrome, hypothyroidism, cerebral palsy, cataracts); cardio
 CC vascular disease (arteriovenous fistula, hypertension, vasculitis,
 CC aneurysms, congestive heart failure, angina pectoris, myocarditis,
 CC ischaemic heart disease, chronic bronchitis, lung tumours); lipid
 CC disorder (fatty liver, Fabry's disease, Niemann-Pick's disease,
 CC hypcholesterolaemia, obesity). PKIN DNA is useful for assessing
 CC toxicity of a test compound and in gene therapy. The present sequence
 CC is human PKIN-17 cDNA.
 XX Sequence 2897 BP; 599 A; 921 C; 877 G; 500 T; 0 other;
 SQ
 Query Match 48.1%; Score 964.4; DB 24; Length 2897;
 Best Local Similarity 69.8%; Pred. No. 2.1e-199;
 Matches 1490; Conservative 0; Mismatches 451; Indels 195; Gaps 6;
 QY 12 GGGGAAGACCGCGCGAGCAGCGCGACGTATGTTGGCCCTTACCGCTGGAGAGAC 71
 DB 105 GGAGCCGAAGAGAGAGCGCGCCACCAATATGTGGGCCCTTATCGCTGGAGAGAC 164
 QY 72 GCTGGGCAAGGGGAGACAGAGGTCTGGTGAAGCTGGGGGTTCACTCGCTCACTGCCAGAA 131
 DB 165 GCTGGGCAAGGAGACAGACAGGCGTGGTTAACTCGGGTCCACTGCGATCATCAGGGTCAGAA 224
 QY 132 GGTGCCATCAAGATCGTCAACCGGTGAGAGCTCAGCGAGTCGGTGTGATGAAGGTGGA 191

DB 225 GGTGCCATCAAGATCGTGAACCGGAGAAAGCTGTCGGAGTCGGTCTGATGAAGGTGGA 284
 QY 192 GCGGAGATCGGATCGTGAAGCTCATTTAGACACCCCGACGTCCTTAAAGCTCGACGACGT 251
 DB 285 GCGGAGATCGGCATCTCTGAAGCTCATCGAACCCACCATGTCCTCAAGCTCCACGACGT 344
 QY 252 TTATGAAAAAATAATATTTGTACTGCTGTAGAACACAGTGTGAGGTGTGAGCTCTT 311
 DB 345 CTACGAGAACAGAAATATTTGTACTGCTGTGAGACAGCTCTCGGGGGGTGAGCTATT 404
 QY 312 CGACTACCTGTGTAAGAGGGAGCTGACGCTTAAGAGGCTCGAAGTCTTTCGGGCA 371
 DB 405 CGACTACCTGTGTAAGAGGGAGCTGACGCCCAAGGAGGCCGAAAGTCTTTCGGCCA 464
 QY 372 GATCATCTCTGCGCTGGACTTCTGCCACAGGCACATCCATATATCCACAGGATCTGAAACC 431
 DB 465 GATGTGTCTGCGCTGGACTTCTGCCACAGCTACTTCCATCTGCCACAGAGACTTAAGCC 524
 QY 432 TGAAGAACCTCTCTGTCGAGAGAGAACAAATCCGCAATCGAGACTTTTGGCATGCGCTC 491
 DB 525 CGAGAACCTCTCTTTTGGATGAGAAAAAACAATCCGCAATTCGAGACTTCGGCATGCGCTC 584
 QY 492 CTTGACAGTTGGGACAGCTCTGTTGAGACAGCTGTGAGTCCGCCCATACGCTGCGCTGCC 551
 DB 585 CTTGACAGTTGGGACAGCTCTCTGAGACAGCTGCGGGTCCGCCCATTTATGCGTGTCC 644
 QY 552 CGAGGTGATCCCGGGGAGAGATATGACGGCGGAGGCGGAGCTGTGAGCTGCGGCGT 611
 DB 645 AGAGGTGATTAAGGGGAAAAATATATGTCGCGCGGACAGATGTGAGCTGTGAGT 704
 QY 612 CATCTCTTTCGCTTGTCTGTTGGGGGCTCTGCTTTCGAGTACAACTTCGACAGCT 671
 DB 705 CATCTCTTTCGCTTGTCTGTTGGGGGCTCTGCTTTCGAGTACAACTTCGCGAGCT 764
 QY 672 GCTGAGAGGTGAGCGGGGCTGTTCCACATCGGCACTTTATCCCGCCGACTGCCA 731
 DB 765 GCTGAGAGGTGAGAAACGGGGCTTTCACATGCCCCCTTCACTTCTCCAGATTGCCA 824
 QY 732 GAGTCTCTACCGGGGATGATCGAGGTGACGCGCGACCGCTCACCTAGAGACAT 791
 DB 825 GAGCTCTCTGAGGGGATGATGAGTGGAGCGCGGAGGCTCAGTCTGGAGCAAT 884
 QY 792 TCAGAAACATATGTTATATAGGGGGCAAGATAGCCCGAAC-----AGAGCAGCC 845
 DB 885 TCAGAAACATCTCTGTGTAAGGCGGGGAAACACAGAGCCAGACCCGCTGCGAGCAGC 944
 QY 846 CATCTCTCCAGAGTGCAGATCGCTGCTGCCAGCTTGGAGGAGCATCGACCCGACGT 905
 DB 945 CCGTGGCCCGGGGTAGCCATGCGAGCTGCGCATCCAGCGAGAGCTGGACCCCGACGT 1004
 QY 906 GCTGAGACAGATGCACTCACTGGGCTGCTTCCGAGACCGCAACAACTGCTGCGAGACCT 965
 DB 1005 CTTAGAGAGCATGCGATCATCTGGCTGCTTCCAGGACCGGAGAGCTGCATCGAGCT 1064
 QY 966 GCTGTCGAGAGGAGAGACAGAGAGATGATTTACTTCTCTCTGAGACCGGAAAGA 1025
 DB 1065 GCGCAGTGGAGGAGAGAACCAAGAAAGATATATTTCTGCTTTTGGATCGGAAGGA 1124
 QY 1026 AAGGTACCCGAGCCAGGAGATGAGGACCTGCCCGCCCGGACGAGATAGACCTCCCG 1085
 DB 1125 GCGGTATCCAGCTGTGAGGACCGAGACCTGCTCCCGGAAATGATGATGACCCCGCCG 1184
 QY 1086 GAAGCGTGTGACTCCCGCATGCTGAACCGGCAACGCGGCGCGGCGGCAAGCAATC 1145
 DB 1185 GAAGCGTGTGATTTCTCCCATGCTGAGCCGCTCACCGGAGGCGGCGGAGGAGTC 1244
 QY 1146 CATGGAGGTGCTCAGCTGAC-----GAGCGGGCTCCCGGTGCTCGCGGCGG 1196
 DB 1245 CATGGAGGTGCTCAGCTGAC-----GAGCGGGGTGGTGGCTTCCCTGTATCCACCGGACG 1304
 QY 1197 GGCATTGAGATGCGCCGAGCGGAGGCTCGGTCCATCAGCGGTGCTCTCAGG 1256

QY 661 TTGCGACAGCTGCTGGAGAAAGGTGAAGCGGGCGTGTTCACATGCGCCACTTTATCCCG 720
 Db 481 CTCCGCCAGCTGCTGGAGAGGTGAAGCGGGCGTGTTCACATGCGCCACTTTATCCCT 540
 QY 721 CCCAGCTGCCAGAGTCTGCTACCGGGGATGATGATGAGGTGAGCGCGCGCGCTCAGC 780
 Db 541 CCAGATTGCCAGAGCTCTCTGAGGGGAATGATCAAGTGGAGCCCGAAAAAGGCTCAGT 600
 QY 781 CTAGAGACACATTCAAGAACACATATGTATATAGGGGGCAAGATGAGCCCGAACCC--- 836
 Db 601 CTGGAGCAAAATTTCAGAAACATCTCTTGTGTACCTAGCGGGGAAACACAGAGCCAGACCCGTGC 660
 QY 837 --AGAGCAGCCCATTTCTTCGAAGGTGCAGATCCGCTCGCTGCCAGCCCTGGAGGACATC 894
 Db 661 CTGGAGCAGCCCTTGGCGCGCGGTAGCCATCGGAGCCTGCCATCAACGGAGAGCTG 720
 QY 895 GACCCGAGCTGCTGGACAGATGCTACTCTGGGCTGCTTCGAGACCGCAACAGCTG 954
 Db 721 GACCCGAGCTCTTAGAGAGCATGGCATCACTGGGCTGCTTCAGGGACCGCGAGAGCTG 780
 QY 955 CTGAGGACCTGCTGCTCGAGGAGGAAACAGAGAGAGATGATTTACTTCTCTCTCTG 1014
 Db 781 CATCGGAGCTGCGAGTGAAGAGAGAACCAAGAAAGATGATATATATCTCTCTTTG 840
 QY 1015 GACCGGAAGAAAGGTATCCGAGCCAGAGGATGAGGACCTGCCCCCGGAAACGAGATA 1074
 Db 841 GATCGGAAGAGCGGTATCCCGAGCTGTGAGGACAGGACCTGCCCTCCCGGAATGATGT 900
 QY 1075 GACCTCTCCCGAAGCTGTGGATCTCCCGATGTCTGAACCGGACGGCAAGCGGGGCCA 1134
 Db 901 GACCCCGCCGGAAGCGTGTGGATTCTCCATGCTGAGCCGTCAAGGAAAGCGCGACCA 960
 QY 1135 GAACCAAAATCCATGAGGTGCTAGCTGTACGGA-----CGGCGGCTCCCGGTG 1185
 Db 961 GAGCGGAAGTCAATGGAAGTCTGAGCATCACGATGCCGGGGTGTGGTCTCCCTGTGTA 1020
 QY 1186 CTTGCGCGCGGGCCATTTAGATGATGCCAGACGCGCCAGAGGTCTCGGTCCATCAGCGGT 1245
 Db 1021 CCCACCCGAGCGGCTTGGAGATGATGCCAGACAGCAGAGATCCCGTAGCGTCAGTGGA 1080
 QY 1246 GCCTCTCAGGCTTTCCACAGCCCATCTAGCAGCCCCCGG----- 1288
 Db 1081 GCCTCCAGGGTCTGTCTCCAGCCCTCTAAGCAGCCCAAGGAGTCCGGTCTTTTCTCTTT 1140
 QY 1289 ----- 1288
 Db 1141 TCACCGGAGCGGGGCTGGAGATGAGGCTCGAGCGGGGGCTCCCGACTTCCAAAACG 1200
 QY 1289 -----TGACCCCTCAGCCCC 1302
 Db 1201 CAGACGCTGCTTCTCGGGGCCAGGAGGTGGGGCGCGCGGAGCAGCCCCCGCCCC 1260
 QY 1303 TCACCAAGGGGAGTCCCTCTCCG-----ACC 1329
 Db 1261 AGTGGCGCTCCACACCCCTGCTCCCGGCCCCCAGGCTCTCCCGCGCTCTCTGCGGGGACC 1320
 QY 1330 CCCAAGGGGACACCTGTCCACACGCAAGAGGAGCGCGGTGGCAGCAGCCCAACCCACG 1389
 Db 1321 CCTTTGCACTCGCTCTGACACAGCCCGGGCCAGTCCCAACGGGACCCCGGGGACACA 1380
 QY 1390 CCCCCTGTCAGCCCC-----AGGTGGAAGGGGTGCTCTGAGGGCGCGGTCAACTCC 1443
 Db 1381 CCACCCCGCAGCCCGGGGCGGTGGCGTCCGGGGAGCGCGCTGAGGAGTCTCTCAACTCC 1440
 QY 1444 ATCAGAGACGCTTTCTGGGCTCACCCGCTTCCACCGCGGAACTCAAGTTCCGAGC 1503
 Db 1441 ATCCGCAACAGCTTCTGGGCTCTCCCTGCTTTTACCGGCGCAAGATGAGGTCTCTACC 1500
 QY 1504 CCGGAGGAGATCTCAACCTGACACAGAGTCTCTCCAGAGCTGGCGAAGAGTCTCTGG 1563
 Db 1501 GCTGAGAGATCTCAGCTTGAAGCAGAGTCTCTCCCGGAGCTGGCAAAACGCTCTCTGG 1560

QY 1564 TTTGGAACTTCATCAGCGCTGGAGAAAGGAGGAGAGATCTTCGTGGTCAATCAAGACAAA 1623
 Db 1561 TTTGGAACTTCATCTCTTGGACAAAGAGAAACAAATATCTCTGTGCTAAAGACAAA 1620
 QY 1624 CTTCTGAGCTCCATCAAGCTGACATCGTGGACGCTTCCTGTCGATTCCTCCAGTCTCAGC 1683
 Db 1621 CTTCTCAGCAGCATCAAGCAGACATCGTCCATGCTTTCTGTGATCCCAAGCTTGA 1680
 QY 1684 CACAGCGTCATCTCCCAACAGAGCTTCCGGGCGGAGTACAAGGCCACGCGGGGGCCAGCC 1743
 Db 1681 CACAGTGTCTGTACAGACAGCTTTCAGGGCGGAGTACAAGGCCAGTGGCGGCCCTCC 1740
 QY 1744 GTGTTCCAGAACCGGTCAAGTTCAGAGTTGATATCACTACACGAGAGGT----- 1794
 Db 1741 GTCTTCAAAAGCGGCTCGCTTCCAGGTGGACATCAGCTCTCTGAGGTCCAGAGCCC 1800
 QY 1795 -----GGGAGGCGCAGAGGAGAACGCGCATCTACTCCGTCACTTCACTTCACTCTC 1845
 Db 1801 TCCCCCGGACGGGACGGCAGCGGAGGTGGTGGCACTACTCCGTCACTTCACTTCACTCTC 1860
 QY 1846 TCAGGCCCCCAGCGCTCGCTTCAAGAGGGTGGTGGAGACCATCCAGGCCCAGCTTGTGAGC 1905
 Db 1861 TCGGGTCCAGCGCTCGCTTCAAGCGAGTGGTGGAGACCATCCAGGCACAGCTCTCTGAGC 1920
 QY 1906 ACACAGACCCCGCTCGGCGCCAGCAGCTTGTCAAGACACCACTAACTG 1952
 Db 1921 ACTCATGACGAGCCCTTCGTGCGGCGCTTGGCAGAGCGGAGAGAACGG 1967

RESULT 13

ABV74558
 ID ABV74558 standard; cDNA; 2289 BP.
 XX AC ABV74558;
 XX DT 20-JAN-2003 (first entry)
 XX DE Human kinase #2 coding sequence.
 XX KW Human; kinase; chromosome 19; gene; ss.
 XX OS Homo sapiens.
 XX FH Key Location/Qualifiers
 XX CDS 1..2289
 XX FT /*tag= a
 XX FT /product= "Human kinase #2".
 XX PN W0200281670-A1.
 XX PD 17-OCT-2002.
 XX PF 04-APR-2002; 2002MO-US10786.
 XX PR 06-APR-2001; 2001US-282036P.
 XX PA (LEXI-) LEXICON GENETICS INC.
 XX PI Turner CA, Mathur B, Friddle CJ;
 XX WPI; 2003-058538/05.
 XX P-PSDB; ABB98744.
 XX PT New human kinase proteins useful for diagnosis, drug screening, and
 XX PT clinical trial monitoring, treatment of disorders and diseases, and
 XX PT cosmetic and nutritional applications -
 XX PS Claim 5; Page 41-42; 47pp; English.
 XX CC The present sequence is the coding sequence for a novel human kinase. The
 XX CC genomic locus encoding the kinase is thought to be on human chromosome
 XX CC 19. The kinase and its coding sequence are useful for diagnosis, drug
 XX CC screening, clinical trial monitoring, treatment of disorders and

XX Human; gene; ss; nervous system disorder; peripheral neuropathy;
 KW Huntington's disease; amyotrophic lateral sclerosis; haemophilia;
 KW neurodegenerative disease; Parkinson's disease; Alzheimer's disease;
 KW autoimmune disease; systemic lupus erythematosus; rheumatoid arthritis;
 KW insulin-dependent diabetes mellitus; anaemia; thrombocytopaenia; wound;
 KW ulcer; burn; bone disorder; osteoporosis; osteoarthritis; stroke;
 KW fibrosis; reperfusion injury; infection; allergic rhinitis; asthma;
 KW coagulation disorder; cancer; tumour; inflammatory disease;
 KW septic shock; Crohn's disease; anaphylaxis; proliferation; chemotactic;
 KW differentiation; stem cell growth factor; haematopoiesis; chemokine;
 KW haemostatic; antiinflammatory; expressed sequence tag; EST.
 XX Homo sapiens.
 XX W0200281731-A2.
 XX 17-OCT-2002.
 XX 29-JAN-2002; 2002WO-US01222.
 XX 30-JAN-2001; 2001US-0774528.
 XX (HYSE-) HYSEQ INC.
 XX (GOOD/) GOODRICH R. W.
 XX Tang TY, Liu C, Zhou P, Asundi V, Zhang J, Zhao QA, Ren F;
 XX Xue AJ, Yang Y, Wehrman T, Wang J, Wang D, Drmanac RT;
 XX WPI; 2003-058563/05.
 XX Novel polypeptide useful for treating neurodegenerative diseases,
 XX myeloid or lymphoid cell disorders, bone disorders, mechanical and
 XX traumatic disorders, coagulation disorders, and inflammatory diseases
 XX .
 XX Claim 1; Page -; 612pp; English.
 XX This invention relates to the cDNA sequences encoding an isolated
 XX novel human polypeptide. The protein encoded by the nucleic acid of
 XX the invention is useful for treating central and peripheral nervous
 XX system diseases (e.g. peripheral neuropathy, Huntington's disease,
 XX amyotrophic lateral sclerosis); neurodegenerative diseases (e.g.
 XX Parkinson's disease, Alzheimer's disease); autoimmune disease (e.g.
 XX systemic lupus erythematosus, rheumatoid arthritis, insulin-dependent
 XX diabetes mellitus); myeloid or lymphoid cell disorders (e.g. anaemia
 XX and thrombocytopaenia); wounds, ulcers, burns; bone disorders (e.g.
 XX osteoporosis, osteoarthritis); mechanical and traumatic disorders (e.g.
 XX stroke, head trauma); lung or liver fibrosis; reperfusion injury in
 XX various tissues; bacterial, viral or fungal infections; allergic
 XX conditions such as allergic rhinitis, asthma; coagulation disorders
 XX (e.g. haemophilia); cancer and tumours; and inflammatory diseases (e.g.
 XX septic shock, Crohn's disease, anaphylaxis). The protein may be used to
 XX inhibit the growth, infection or function of infectious agents such as
 XX bacteria, fungi, viruses, or to effect bodily characteristics,
 XX biorhythms or circadian cycles of rhythms. The protein may also
 XX have proliferation/differentiation, stem cell growth factor,
 XX haematopoiesis regulation, immune stimulating or suppressing,
 XX chemotactic/chemokinetic, haemostatic and thrombolytic, receptor/ligand,
 XX and antiinflammatory activities. The cDNA sequences of the invention are
 XX useful for expressing recombinant protein for analysis. The present
 XX sequence represents a novel human cDNA sequence of the invention,
 XX this sequence is an expressed sequence tag (EST) and was identified
 XX using subtractive hybridisation.

XX Sequence 1797 BP; 375 A; 549 C; 520 G; 353 T; 0 other;

Query Match 29.3%; Score 588.8; DB 25; Length 1797;
 Best Local Similarity 77.1%; Pred. No. 5.2e-118;
 Matches 748; Conservative 0; Mismatches 207; Indels 15; Gaps 2;

QY 273 GTACCTGGTGTAGAACAGCTGCTAGGTGGTGGTCTTCTGACTACTCTGGTGAAGAGGG 332
 |||||

Db 89 GTACCTGGTCTTGGAGCAGCTCTCGGGGGGTGAGCTATTTCGACTACTCTGTTAAAGAGGG 148
 QY 333 GAGGCTGACGCTTAAGAGAGCTCGGAAGTCTTCGGGAGATCATCTCTGCGCTGACTT 392
 Db 149 GAGACTGACGCGCCAGAGAGCGCCGAAAGTCTTCGCCAGATTTGTCTGCGCTGACTT 208
 QY 393 CTGCCACAGCAGCTTCCATATGCCCACAGGGATCTGAAACCTGAAACCTCTCTGCTGGACGA 452
 Db 209 CTGCCACAGCTACTCCATCTGCCACAGAGACCTTAAAGCCGAGAACCTGCTTTGGATGA 268
 QY 453 GAAGAACACATCCGATCGAGACTTTGGCATGGCGTCCCTGCGAGTTGGCGACAGCCT 512
 Db 269 GAAAAACAACATCCGATTCGAGACTTCGGCATGGCGTCCCTGCGAGTTGGCGAGCCT 328
 QY 513 GTTGGAGAGCAGCTGTGGGTCCCGCCACTACGCTCCCGAGGTGATCCGGGGGAGAA 572
 Db 329 CTGGAGAGCAGCTGCGGTCCCGCCACTTATGGGTCTCAGAGGTGATTAAGGGGAGAA 388
 QY 573 GTATGACGCGCGAAGCGGACGCTGTGGAGCTGCGCGCTCATCTCTTTCGCTTGTGCT 632
 Db 389 ATATGATGCGCGCGGAGACATGTGGAGCTGTGGAGTCTCTTTCGCTTGTGCT 448
 QY 633 GGGGCTCTGCGCTTCGATGACATGACAACTTGGACAGCTGCTGGAGAGAGGTGAAGCGGG 692
 Db 449 GGGGCTCTGCGCTTTCATGACGACAACTTCGCCAGCTGCTGGAGAGAGGTGAAGCGGG 508
 QY 693 CGTGTTCACATGCGCGACTTTATCCGCGCGAGTCCGAGAGTCTGCTACGGGGCATGAT 752
 Db 509 CGTCTCCACATGCGCGACTTTATCTCCAGATTGCGAGGCTCTCTGAGGGGATGAT 568
 QY 753 CGAGGTGGAGCGCGCGCTCAGCTAGAGCATTTCAGAAAACATATGGTATAT 812
 Db 569 CGAAGTGGAGCGCGGAAAAAGGCTCAGTCTGGAGCAATTCAGAAACATCTCTTGGTACCT 628
 QY 813 AGGGGCGAGAGTACGCGCGAACCA-----GAGCAGCCCATTCCTCGAGGTGCGAT 866
 Db 629 AGCGGGGAAACACGAGCCAGACCCCTGCTGGAGCCAGCCCTGGCGCGGGTAGCCAT 688
 QY 867 CCCTCGCTGCGCGAGCTGAGGAGCATCGACCCCGAGCTGCTGGAGAGCATGCACTCACT 926
 Db 689 CGGAGCTGCTCCATCCACGAGAGCTGACCCCGAGCTCTAGAGAGCATGCGATCACT 748
 QY 927 GGGCTGCTTCGAGAGCGCGAACAGCTGTGAGGAGCTGCTGCTCGAGAGAGGAGAACCA 986
 Db 749 GGGCTGCTTCAGGAGCGCGAGAGGCTGCAATCGGAGCTGCGCAGTGCAGAGGAGAACCA 808
 QY 987 GAGAGAGTATTACTTCTCTCTCGACCGGAAAGAGGTACCCGAGCCAGAGGA 1046
 Db 809 AGAAAAGATATATATATATCTGCTTTTGGATCGGAAGAGCGGTATCCGAGCTGAGGA 868
 QY 1047 TGAGGACCTGCGCGCGGAAACAGATAGACCTCCCGGAGCGGTGTGGACTCCCGAT 1106
 Db 869 CGAGGACCTGCTTCCCGGAGTATGATTTGACCCCGCGAGAGCGGTGTGGATTTCCCAT 928
 QY 1107 GCTGAACCGGACCGGAGCGCGCGAGAACGCAATTCATGAGAGGTGCTCAGCGTGAC 1166
 Db 929 GCTGAGCGCTCAGCGGAGCGCGGACGAGAGCGGAGTCCATGGAAGTCTCTGAGCATCAC 988
 QY 1167 -----GAGCGCGCTCCCGGCTGCTGCGGGGGGCGCATTTGAGATGCCCCAGCA 1217
 Db 989 CGATGCGGGGGGTGGTGGTCTCCCTGTACCCACCGGAGCGGCTTGGAGATGGCCCCAGCA 1048
 QY 1218 CGGCGAGAGG 1227
 Db 1049 CAGCCAGAGG 1058

RESULT 15
 ABL10489
 ID ABL10489 standard; cDNA; 2720 BP.
 XX
 AC ABL10489;
 XX

DT 26-MAR-2002 (first entry)
XX Drosophila melanogaster expressed polynucleotide SEQ ID NO 25949.
DE XX
KW Drosophila; developmental biology; cell signalling; insecticide;
KW pharmaceutical; gene; ss.
XX Drosophila melanogaster.
OS
XX WO200171042-A2.
XX 27-SEP-2001.
XX 23-MAR-2001; 2001WO-US09231.
XX 23-MAR-2000; 2000US-191637P.
XX 11-JUL-2000; 2000US-0614150.
XX (PEKE) PE CORP NY.
XX Venter JC, Adams M, Li PWD, Myers EW;
PI WPI; 2001-656960/75.
XX P-PSDB; AB566386.
XX New isolated nucleic acid detection reagent for detecting 1000 or more
PT genes from Drosophila and for elucidating cell signalling and cell-cell
PT interactions -
XX Claim 1; SEQ ID NO 25949; 21pp + Sequence Listing; English.
XX The invention relates to an isolated nucleic acid detection reagent
XX capable of detecting 1000 or more genes from Drosophila. The invention is
XX useful in developmental biology and in elucidating cell signalling and
XX cell-cell interactions in higher eukaryotes for the development of
XX insecticides, therapeutics and pharmaceutical drugs. The invention
XX discloses genomic DNA sequences (AB16176-AB130511), expressed DNA
XX sequences (AB101840-AB16175) and the encoded proteins
XX (AB57737-AB572072).
XX The sequence data for this patent did not form part of the printed
XX specification, but was obtained in electronic format directly from WIPO
XX at ftp.wipo.int/pub/published_pct_sequences.
XX Sequence 2720 BP; 661 A; 715 C; 759 G; 585 T; 0 other;
Query Match 26.9%; Score 539.8; DB 23; Length 2720;
Best Local Similarity 70.4%; Pred. No. 2.4e-107;
Matches 737; Conservative 0; Mismatches 307; Indels 3; Gaps 1;
QY 9 GACGGGGAAGGACGGGGCGCGCAGCACCGCAGTATGTTGGGCGCTACCGCTGGAGAA 68
DB 6 GAAGGAGAAACAATGTACCGCGGAGAAATTCGCAATTTGTGGGGCGCTATCGCTGGAGAA 65
QY 69 GACGCTGGGCAAGGGGACACAGGCTCTGCTGAAGCTGGGGGTTCACTGGTCCACCTGCCA 128
DB 66 AACCTTGGCAAGGGTCAACAGGGTCTGCTCAAGTTGGGGCGTGCATTTGTGATTTGGCAA 125
QY 129 GAAGGTGGCCATCAAGATCGTCAACCGTGAAGAGCTCAGCAGTGGGTGCTGATGAAGT 188
DB 126 GAAGGTGGGATTAATAATCAATCGCGAGAACTCACGGAATCGGTCTAATGAAGT 185
QY 189 GGAGCGGGAGATCGGATCCTGAAGCTCATTTGAGCACCCCGTCTTAAGCTGCA 248
DB 186 TGAACGTGAAATCGGCATTAATAATCGATCATCCACACGCTCTTGGCCTGAGCGA 245
QY 249 CGTTTATGAAACAAATAATTTGTACCTGGTGGTGTAGAACACGCTGTGAGGTGGT 308
DB 246 TGTGTACGAGAACAGAAATTTTGTATTTATTTGAGCATGTATCCGGCGGAGAGCT 305
QY 309 CTTCCGACTACCTGGTGAAGAGGGAGGCTGACGCTTAAGAGGCTCGGAAGTTCTTCG 368
DB 306 CTTCCGATTACCTGGTGAAGAGGGTGGATTGACCGCGAAGGAGGCGCGCAAGTTCTTCAG 365

QY 369 GCAGATCATCTCTGGCGTGGACTTTCGCCACAGCCACTCCATATGCCACAGGATCTGAA 428
DB 366 GCAATCATCTCGCGCCCTGGATTTCTGCCACTCGCATTCGATTTGCCATCGCGACTTGA 425
QY 429 ACCTGAAAACTCTCTGCTGGACGAGAGAAACAACATCCGATCCGACACTTTGCGATGGC 488
DB 426 GCCGAGAGATCTGCTGTGACGAGAGAAATAACATTAAGNATAGCGACTTTGGATGGC 485
QY 489 GTCCCTCAGGTTGGCGACAGCCCTTTGGAGACAGCTGTGGGTCCCCCACTACGCGCTG 548
DB 486 TTCCCTCAGCCAGCTGGCAGCATGTTGGAGACCTCTCTCGCGGACCCACACTACCGGTG 545
QY 549 CCCGAGGTGATCCGGGGGAGAGTATGACGGCCGGAAGCGGACGCTGTGGAGCTGGG 608
DB 546 TCCAGAGGTATACCGGGGCGAGAAAGTACGATGGCCGCAAGCGGATGCTCTGGTCTG 605
QY 609 CGTCATCTGTTTCGCTTGTGGTGGGGCTCTCCCTTCGACGATGACAACTTTGCCACA 668
DB 606 GGTATCTCTATGCCCTCTCTGTTGGTGGGTGCGTTCCTTCGACGAGCAACTTTGGCGCA 665
QY 669 GCTCTGGAGAGGTGAAGCGGGCGTGTTCACATGCCGCACTTTATCCCGCGGACTG 728
DB 666 GCTCTGGAGAAAGTCAAGCGGGCGTCTTTTACATACCGCACTTTGTGCGCGGACTG 725
QY 729 CCAGACTCTGACGGGGCATGATCGAGGTGGACCGCGCAGCGCCCTCAGCTAGAGCA 788
DB 726 CCRGAGTCTGCTGCGCGCATGATTTGAGGTCAATCCCGACCGCGGCTCACGCTGGCTGA 785
QY 789 CATTGAGAAACACATATGGTATATA---GGGGGCAAGAAATGAGCCCGAACCCAGAGCAGCC 845
DB 786 AATCAACCGGATCCGTGGGTCAAGCTGGCGGCAAGGGAGCTGGAGCTGGAGCTGCC 845
QY 846 CATTCTCGCAAGGTGCGAGATCCGCTCGTCCGAGCTGGAGGACATCGACCCGAGCT 905
DB 846 AATGATGGAGGTGTGCGACACACACGTTATTTCCACACAGCCACCGCGGTGGATCCGGATGT 905
QY 906 GCTGGACAGCATGCACTCAGTGGGCTGCTTCGAGACCGCAACAGCTGCTGCGAGGACT 965
DB 906 GTTGAACGCGATTTGCTGCTGGGCTGTTTCAAGGAGAGGAGAACTCATCCAGGAACT 965
QY 966 GCTGTCGAGGAGAGAACACGAGGAGAGATGATTTACTTCTCTCTGAGACCGGAAAGA 1025
DB 966 GCTCAGTTCAAGTCACATACGAGAGAGGTATATATTTCTCTGTTGCTCGAGCGCAACG 1025
QY 1026 AAGGTACCCGAGCCAGGAGGATGAGGA 1052
DB 1026 AAGACGACCTCGCTGGAGGATGATGA 1052

Search completed: November 27, 2003, 05:14:47
Job time : 524.488 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 27, 2003, 02:00:45 ; Search time 7183.11 Seconds
(without alignments)
11430.368 Million cell updates/sec

Title: US-10-054-579-1
Perfect score: 2007
Sequence: 1 atgacatcgacggggaagga.....gaattatccgaaaagtaa 2007

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 2888711 seqs, 2045481386 residues

Total number of hits satisfying chosen parameters: 5777422

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : GenEmbl.*
1: gb.ba.*
2: gb.htg.*
3: gb.in.*
4: gb.om.*
5: gb.ov.*
6: gb.pat.*
7: gb.ph.*
8: gb.pl.*
9: gb.pr.*
10: gb.to.*
11: gb.ats.*
12: gb.sy.*
13: gb.un.*
14: gb.vi.*
15: em.ba.*
16: em.fun.*
17: em.hum.*
18: em.in.*
19: em.mu.*
20: em.om.*
21: em.or.*
22: em.ov.*
23: em.pat.*
24: em.ph.*
25: em.pl.*
26: em.ro.*
27: em.sts.*
28: em.un.*
29: em.vi.*
30: em.htg.hum.*
31: em.htg.inv.*
32: em.htg.other.*
33: em.htg.mus.*
34: em.htg.pln.*
35: em.htg.rtd.*
36: em.htg.mam.*
37: em.htg.vrt.*
38: em.sy.*
39: em.htgo.hum.*
40: em.htgo.mus.*
41: em.htgo.other.*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	2007	100.0	2007	6	AX661191	AX661191 Sequence
2	2005.4	99.9	3516	6	AR232170	AR232170 Sequence
3	2005.4	99.9	3516	6	AY168857	AY168857 Homo sapi
4	1993	99.3	3364	6	AR232171	AR232171 Sequence
5	1939.6	96.6	2025	6	AX327995	AX327995 Sequence
6	1939.6	96.6	2219	6	AX327993	AX327993 Sequence
7	1827	91.0	1827	6	AX661193	AX661193 Sequence
8	1727	86.0	1956	9	HSAG701	AJ006701 Homo sapi
9	1330.4	66.3	3156	9	AK074411	AK074411 Homo sapi
10	965	48.1	3109	9	AF479827	AF479827 Homo sapi
11	964.4	48.1	2385	6	AX166526	AX166526 Sequence
12	964.4	48.1	2897	6	AX642966	AX642966 Sequence
13	964.4	48.1	3007	9	AF479826	AF479826 Homo sapi
14	946	47.1	2720	9	HSAG05307	AL834275 Homo sapi
15	878.8	43.8	2576	9	AB058714	AB058714 Homo sapi
16	594	29.6	2333	9	AF020089	BC024291 Homo sapi
17	583.2	29.1	1873	9	BC024291	BC024291 Homo sapi
18	432.2	21.5	3933	3	AB014885	AB014885 Halocynthia
19	426.6	21.3	3059	3	AF316542	AF316542 Caenorhab
20	426	21.2	5609	3	AY060288	AY060288 Drosophila
21	410.2	20.4	2123	3	AK116009	AK116009 Ciona int
22	391	19.5	2128	9	BC016681	BC016681 Homo sapi
23	265.8	13.2	3529	9	AB088047	AB088047 Homo sapi
24	265.8	13.2	3609	9	AB049127	AB049127 Homo sapi
25	265.6	13.2	2259	10	AY151083	AY151083 Mus muscu
26	265	13.2	1594	6	AX056390	AX056390 Sequence
27	264.2	13.2	1549	6	AX685979	AX685979 Sequence
28	264.2	13.2	2462	6	AX399987	AX399987 Sequence
29	264.2	13.2	2650	9	AY120867	AY120867 Homo sapi
30	264.2	13.2	2954	6	BD127952	BD127952 Primer fo
31	264.2	13.2	2954	9	AK075272	AK075272 Homo sapi
32	264.2	13.2	3226	9	AY057448	AY057448 Homo sapi
33	264.2	13.2	3269	6	AX680149	AX680149 Sequence
34	264.2	13.2	3312	6	AX305105	AX305105 Sequence
35	264.2	13.2	3392	6	AX305106	AX305106 Sequence
36	264.2	13.2	4917	9	AB058763	AB058763 Homo sapi
37	262.4	13.1	3170	6	AX305103	AX305103 Sequence
38	262.4	13.1	3250	6	AX305104	AX305104 Sequence
39	259.8	12.9	1749	3	AF020310	AF020310 Drosophila
40	257.2	12.8	2650	10	AB020480	AB020480 Rattus no
41	257.2	12.8	4092	10	AF106937	AF106937 Rattus no
42	255	12.7	1791	9	AB062946	AB062946 Macaca fa
43	252.4	12.6	2352	6	AX262512	AX262512 Sequence
44	252.4	12.6	2968	6	AX262510	AX262510 Sequence
45	252.4	12.6	4699	6	AX224729	AX224729 Sequence

ALIGNMENTS

RESULT 1
AX661191
LOCUS AX661191
DEFINITION Sequence 1 from Patent WO02059287.
ACCESSION AX661191
VERSION AX661191.1 GI:29162844
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
AUTHORS Turner, C.A. and Mathur, B.
TITLE Novel human kinases and polynucleotides encoding the same
JOURNAL Patent: WO 02059287-A 1 01-AUG-2002;

2007 bp DNA linear PAT 22-MAR-2003

FEATURES		Lexicon Genetics Incorporated (US)		Location/Qualifiers		1. .2007		/organism="Homo sapiens"		/mol_type="genomic DNA"		/db_xref="taxon:9606"		445 a -626 c 604 g 332 t		BASE COUNT		ORIGIN	
Query Match		100.0%; Score 2007; DB 6; Length 2007;		Best Local Similarity 100.0%; Pred. No. 0;		Matches 2007; Conservative 0; Mismatches 0; Indels 0; Gaps 0;													
QY	1	ATGACATCGACGGGAAGGACGGCGCGCAGCAGCGCAGTATGTTGGCCCTACCGG	60																
DB	1	ATGACATCGACGGGAAGGACGGCGCGCAGCAGCGCAGTATGTTGGCCCTACCGG	60																
QY	61	CTGAGAAAGACGCTGGGCAAGGGGACAGAGGCTGTGTAAGCTGGGGTTCACTCGTC	120																
DB	61	CTGAGAAAGACGCTGGGCAAGGGGACAGAGGCTGTGTAAGCTGGGGTTCACTCGTC	120																
QY	121	ACCTGCCAGAGGTGGCCATCAGATCGTCAACGCTGAGAGCTCAGCGAGTCGGTCTG	180																
DB	121	ACCTGCCAGAGGTGGCCATCAGATCGTCAACGCTGAGAGCTCAGCGAGTCGGTCTG	180																
QY	181	ATGAAGGTGGAGCGGAGATCGCGATCCTGAAGCTCATTGAGCAACCCCACTCTAAAG	240																
DB	181	ATGAAGGTGGAGCGGAGATCGCGATCCTGAAGCTCATTGAGCAACCCCACTCTAAAG	240																
QY	241	CTGCAAGCTTTATGAAAACAAAAATATTTGTAACCTGTGTGCTAGAACATGTCAGGT	300																
DB	241	CTGCAAGCTTTATGAAAACAAAAATATTTGTAACCTGTGTGCTAGAACATGTCAGGT	300																
QY	301	GGTGAAGCTTCGACTACCTGCTGAAGAGGGAGGCTGACGCTAAGAGGCTCGGAAG	360																
DB	301	GGTGAAGCTTCGACTACCTGCTGAAGAGGGAGGCTGACGCTAAGAGGCTCGGAAG	360																
QY	361	TTCTTCGGCAGATCATCTCTGCGCTGGAATTTCTGCCACAGCCACTCCATATGCCAGG	420																
DB	361	TTCTTCGGCAGATCATCTCTGCGCTGGAATTTCTGCCACAGCCACTCCATATGCCAGG	420																
QY	421	GATCTGAACCTGAAACCTCTCTGCTGGAAGAGAAACAAATCCGCGATCGGAGCTTT	480																
DB	421	GATCTGAACCTGAAACCTCTCTGCTGGAAGAGAAACAAATCCGCGATCGGAGCTTT	480																
QY	481	GGCATGGCTCCCTGCAAGTTGGGACAGCTGTTGGAGCCAGCTCTGSGTCCCCCAC	540																
DB	481	GGCATGGCTCCCTGCAAGTTGGGACAGCTGTTGGAGCCAGCTCTGSGTCCCCCAC	540																
QY	541	TACGCTGCCCCGAGGTGATCCGGGGGAGAAATGATGACGGCCGGAAAGCGGAGCTGG	600																
DB	541	TACGCTGCCCCGAGGTGATCCGGGGGAGAAATGATGACGGCCGGAAAGCGGAGCTGG	600																
QY	601	AGCTGGCGCTCATCTGCTTGCCTTGTGTGGGGCTCTGCCCTTCGACGATGACAAC	660																
DB	601	AGCTGGCGCTCATCTGCTTGCCTTGTGTGGGGCTCTGCCCTTCGACGATGACAAC	660																
QY	661	TTGGCAGAGCTGCTGGAGAGGTGAAGCGGGCGTGTTCACATGCCGCACTTTATCCCG	720																
DB	661	TTGGCAGAGCTGCTGGAGAGGTGAAGCGGGCGTGTTCACATGCCGCACTTTATCCCG	720																
QY	721	CCGAGTCCAGAGTCTGCTACGGGGCATGATCGAGGTGAGCGCGCATCGCTTCAAG	780																
DB	721	CCGAGTCCAGAGTCTGCTACGGGGCATGATCGAGGTGAGCGCGCATCGCTTCAAG	780																
QY	781	CTAGAGCATTCAGAAACATATGTTATAGGGGCAAGATGAGCCCGAACACAG	840																
DB	781	CTAGAGCATTCAGAAACATATGTTATAGGGGCAAGATGAGCCCGAACACAG	840																
QY	841	CAGCCATTCCTCGCAAGGTGAGATCGCTCGCTGCCAGCCTGGAGGACATCGACCC	900																
DB	841	CAGCCATTCCTCGCAAGGTGAGATCGCTCGCTGCCAGCCTGGAGGACATCGACCC	900																

Db 1981 AATGTGGAATTATCCGAAAGTTAA 2007
|||||
RESULT 2
AR232170 2908 bp DNA linear PAT 20-DEC-2002
LOCUS
DEFINITION
Sequence 1 from patent US 6455292.
ACCESSION
AR232170
VERSION
AR232170.1 GI:27274061
KEYWORDS
Unknown.
SOURCE
Unknown.
ORGANISM
Unclassified.
REFERENCE
1 (bases 1 to 2908)
AUTHORS
Shu, Y., Fan, W., Kovacs, K.F., Zidanic, M. and Jay, G.
TITLE
Full-length serine protein kinase in brain and pancreas
JOURNAL
Patent: US 6455292-A 1 24-SEP-2002;
FEATURES
Location/Qualifiers
1..2908
/organism="unknown"
BASE COUNT 603 a 948 c 879 g 478 t
ORIGIN
Query Match 99.9%; Score 2005.4; DB 6; Length 2908;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2006; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 ATGACATCAGCGGGGAAGACCGCGCGGCGCAGCAGCGCATGATGTTGGGCGCTACCGG 60
Db 106 ATGACATCAGCGGGGAAGACCGCGCGGCGCAGCAGCGCATGATGTTGGGCGCTACCGG 165
QY 61 CTGGAGAAGACGCTGGGCAAGGGGACAGACAGGTCTGGTGAAGCTGGGGGTTCACTGCGTC 120
Db 166 CTGGAGAAGACGCTGGGCAAGGGGACAGACAGGTCTGGTGAAGCTGGGGGTTCACTGCGTC 225
QY 121 ACTGTCAGAAAGTGCCCATCAAGATCGTCAACCGGTGAGAAGCTCAGCGAGTCCGTGCTG 180
Db 226 ACCTGTCAGAAAGTGCGCATCGGATCCCTGAAGCTATTGAGCAGCCCGACGCTCTAAAG 240
QY 181 ATGAAGGTGGCGGAGATCGGATCCTGAAGCTATTGAGCAGCCCGACGCTCTAAAG 240
Db 286 ATGAAGGTGGCGGAGATCGGATCCCTGAAGCTATTGAGCAGCCCGACGCTCTAAAG 345
QY 241 CTGCAAGCTTTATGAAACAAATAATTTGTACTGCTGTGTAGAACACGCTGTGAGGT 300
Db 346 CTGCAAGCTTTATGAAACAAATAATTTGTACTGCTGTGTAGAACACGCTGTGAGGT 405
QY 301 GGTGAGCTTTGCACTA CTGTTGTAAGAAAGGGAGGCTGACGCTTAAGGAGCTCGGAAG 360
Db 406 GGTGAGCTTTGCACTA CTGTTGTAAGAAAGGGAGGCTGACGCTTAAGGAGGCTCGGAAG 465
QY 361 TTCTTCGGGAGATCATCTCTGGCTGGACTTCTGCGCAGAGCATCTCCATATGCCACAGG 420
Db 466 TTCTTCGGGAGATCATCTCTGGCTGGACTTCTGCGCAGAGCATCTCCATATGCCACAGG 525
QY 421 GATCTGAAACCTTGAAACCTCTGCTGGACGAGAACAACATCCGCATCGCAGACTTT 480
Db 526 GATCTGAAACCTTGAAACCTCTGCTGGACGAGAACAACATCCGCATCGCAGACTTT 585
QY 481 GCGATGGCTCTCTGAGGTTGGCGACAGCTTTGGAGACAGAGTGGGTCCTCCCGCAC 540
Db 586 GCGATGGCTCTCTGAGGTTGGCGACAGCTTTGGAGACAGAGTGGGTCCTCCCGCAC 645
QY 541 TAGGCTGCCCCAGGTGATCGGGGGGAGATGACGGCGGAGAGCGGAGCGACGTGTGG 600
Db 646 TAGGCTGCCCCAGGTGATCGGGGGGAGATGACGGCGGAGAGCGGAGCGACGTGTGG 705
QY 601 AGCTGGGGGTTCATCTGCTGCTGCTGGTGGGGGCTCTGCCCTTCGACCATGACAC 660
Db 706 AGCTGGGGGTTCATCTGCTGCTGCTGGTGGGGGCTCTGCCCTTCGACCATGACAC 765
QY 661 TTGCGACAGCTGCTGGAGAAGGTGAAGCGGGGCGTGTTCACATGCCGCGACTTTATCCCG 720

766 TTGCGACAGCTGCTGGAGAGGTGAAGCGGGGCGTGTTCACATGCGCGCACTTTATCCCG 825
QY 721 CCCGACTGCCAGAGTCTGTACTAGGGGCGCATGATCAGAGTGGAGCGCGCAGCGCTCAGG 780
Db 826 CCCGACTGCCAGAGTCTGTACTAGGGGCGCATGATCAGAGTGGAGCGCGCAGCGCTCAGG 885
QY 781 CTAGAGCACATTTCAAGAACACATATGTTATAGGGGCAAGAAATGAGCCCGAACACAGAG 840
Db 886 CTAGAGCACATTTCAAGAACACATATGTTATAGGGGCAAGAAATGAGCCCGAACACAGAG 945
QY 841 CAGCCCATTTCTCGCAAGGTGCAGATCCGCTGCTGCTGCGCAGCTTGAGAGCATCGACCCC 900
Db 946 CAGCCCATTTCTCGCAAGGTGCAGATCCGCTGCTGCTGCGCAGCTTGAGAGCATCGACCCC 1005
QY 901 GACGTGCTGGACAGCATGCACTCACTGGCTGCTTCGAGACCGCCACAAAGCTGCTGAG 960
Db 1006 GACGTGCTGGACAGCATGCACTCACTGGCTGCTTCGAGACCGCCCAACAGCTGCTGAG 1065
QY 961 GACGTGCTGCTCGAGGAGGAGAACAGAGAGAGATGATTTACTTCTCTCTCTGAGCCGG 1020
Db 1066 GACGTGCTGCTCGAGGAGGAGAACAGAGAGAGATGATTTACTTCTCTCTCTGAGCCGG 1125
QY 1021 AAAGAAAGGTACCCGAGCCAGGAGGATGAGGACCTGCCCGCCCGGAAACGAGATAGACCT 1080
Db 1126 AAAGAAAGGTACCCGAGCCAGGAGGATGAGGACCTGCCCGCCCGGAAACGAGATAGACCT 1185
QY 1081 CCCCGAAGCGTGTGGAATCCCGCATGCTGAACCGGCAAGCGGCAAGCGGCGGCAAGCGC 1140
Db 1186 CCCCGAAGCGTGTGGAATCCCGCATGCTGAACCGGCAAGCGGCAAGCGGCGGCAAGCGC 1245
QY 1141 AAATCCATGAGAGTGTCTCAGCGTGAAGCGGCGGCTCCCGGCTGCTGCGCGGCGGCCC 1200
Db 1246 AAGTCCATGAGAGTGTCTCAGCGTGAAGCGGCGGCTCCCGGCTGCTGCGCGGCGGCCC 1305
QY 1201 ATTGAGATGGCCCCAGCACGCGCAGAGTCTCGGTCCATCAGCGGTGCTCTCTCAGGCTT 1260
Db 1306 ATTGAGATGGCCCCAGCACGCGCAGAGTCTCGGTCCATCAGCGGTGCTCTCTCAGGCTT 1365
QY 1261 TCCACCCAGCCACTCTCAGGAGCGCCCGGGTGACCCCTCACCCCTCACGAGGGGCGAGTCCC 1320
Db 1366 TCCACCCAGCCACTCTCAGGAGCGCCCGGGTGACCCCTCACCCCTCACGAGGGGCGAGTCCC 1425
QY 1321 CTCCCCACCCCAAGGGGAGACCTGTCCACACGCGCAAGAGAGAGCCCGGCTGGCACGCCC 1380
Db 1426 CTCCCCACCCCAAGGGGAGACCTGTCCACACGCGCAAGAGAGAGCCCGGCTGGCACGCCC 1485
QY 1381 AACCCACGCCCCCGTCCAGCCCCAGCGTCCGAGGGGTGCTTCCAGAGGGCGCGCTCAAC 1440
Db 1486 AACCCACGCCCCCGTCCAGCCCCAGCGTCCGAGGGGTGCTTCCAGAGGGCGCGCTCAAC 1545
QY 1441 TCCATCAGAGACAGTCTTCTGGGCTCACCCCGCTTCCACCGCGGGAACCTGCAAGTTCCG 1500
Db 1546 TCCATCAGAGACAGTCTTCTGGGCTCACCCCGCTTCCACCGCGGGAACCTGCAAGTTCCG 1605
QY 1501 ACGCGGAGGAGATGTCACAACTGACACGAGTCTGCCAGAGCTGGCGGAGAGAGTCC 1560
Db 1606 ACGCGGAGGAGATGTCACAACTGACACGAGTCTGCCAGAGTGGTCCGAGAGAGAGTCC 1665
QY 1561 TGGTTTGGGAACTTTCATCAGCTTCGAGAGGAGGAGGAGATCTTCTGTTGGTTCATCAAAGAC 1620
Db 1666 TGGTTTGGGAACTTTCATCAGCTTCGAGAGGAGGAGGAGATCTTCTGTTGGTTCATCAAAGAC 1725
QY 1621 AAACTCTGAGCTCCATCAGGCTGACATCGTGACAGCTTCTCTGATGTTCCAGTCTC 1680
Db 1726 AAACTCTGAGCTCCATCAGGCTGACATCGTGACAGCTTCTCTGATGTTCCAGTCTC 1785
QY 1681 AGCCACAGCTGCTCTCCCAAAACGAGCTTCCGGGCGGAGTACAAGGCCACCGGGGGGCA 1740
Db 1786 AGCCACAGCTGCTCTCCCAAAACGAGCTTCCGGGCGGAGTACAAGGCCACCGGGGGGCA 1845
QY 1741 GCGGTGTTCCAGAAAGCGCGTCAAGTTCAGGTTGATATCACCTTACAGCGAGGTGGGGAG 1800

Db 1846 GCCGTGTTCCAGAGCCGGTCAAGTTCACGTTTCATATACCTACACGAGGGTGGGGAG 1905
Qy 1801 GCGCAGAGGAGAGCGGATCTACTTCGTACCTTACCTTACCTTCTCAGGCCCGCCGCGT 1860
Db 1906 GCGCAGAGGAGAGCGGATCTACTTCGTACCTTACCTTCTCAGGCCCGCCGCGT 1965
Qy 1861 CGCTTCAAGAGGGTGGGAGACCATCAGGCCCGAGCTGTGAGCACACACAGCCCGCT 1920
Db 1966 CGCTTCAAGAGGGTGGGAGACCATCAGGCCCGAGCTGTGAGCACACACAGCCCGCT 2025
Qy 1921 GCGGCCCGAGCTTGTGAGCACACCACTAATGTATGGAATATGATGAGCGGGCGGCTTTCC 1980
Db 2026 GCGGCCCGAGCTTGTGAGCACACCACTAATGTATGGAATATGATGAGCGGGCGGCTTTCC 2085
Qy 1981 AAATGTGAATTATCCGAAAGTTAA 2007
Db 2086 AAATGTGAATTATCCGAAAGTTAA 2112

RESULT 3
AY16857 3516 bp mRNA linear PRI 24-DEC-2002
LOCUS Homo sapiens serine/threonine protein kinase isoform mRNA, complete cds.
DEFINITION
ACCESSION AY16857
VERSION AY16857.1 GI:27369414
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 3516)
AUTHORS Guo, J. H. and Yu, L.
TITLE Direct Submission
JOURNAL Submitted (19-OCT-2002) School of Life Sciences, Fudan University, Institute of Genetics, Handan Rd, 220, Shanghai 200433, China
FEATURES
source
1..3516
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/chromosome="11"
/tissue_type="brain tumor"
387..2393
/note="similar to CAD20436.1"
/codon_start=1
/evidence=not_experimental
/product="serine/threonine protein kinase isoform"
/protein_id="AA087839.1"
/db_xref="GI:27369415"
/translation="WTSQKDGGAQYVGPYRLKLGKGTGLVKLGVHCVTCOK
VALKIVREKLSVLMKVEREIALLLIEHPVHLKLDVYENKKYLVLVLEHVS
LFDILVKGRTLPKARFPKQIISALDFCHSHICHRLAPENLLDEKNRIAPD
GMAIQVDSLELSCGSPKQEVINGEKYDKGKADVWSCGVILFALLVGLPFD
DNLQQLKVRGVFPMHPFIIPDCQSLRGMSEVDRAARLTLEHIOKHIWYIGKNE
PEPEOIPRKVQIRSLPLEDIDPDVLSHSLGCFDRNKLLOLLSEENQEKMY
FLLDKRYPSQEDIDPPRNEIDPPKRVDSPLNHRKRRPERKSMELSVYTDG
SPVPARRAIEAQQHORSISGASSGLSTPLSPRTVPHSPRGSPLTPKCTPVH
TPKSPAGTPNPFPSSVSGVPWRARLINSIKSFLGSPFRHKLQVTPPEMSNL
TPESPALAKSWFNFISLEKEQIFVVIKDKPLSSIKADIVHAFUSIPLSHVSIS
QTSFRAEYKATGCGPAVFPQKVDITTEGGAQKENGIIYSYVTFLLSGPSRRFR
VVEITQALLSTHDPAAQHLSLDTTNCMMMTGRLSKGIIIPKS"

BASE COUNT 642 a 1135 c 1145 g 594 t
ORIGIN

Query Match 99.9%; Score 2005.4; DB 9; Length 3516;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2006; Conservative 0; Mismatch 1; Indels 0; Gaps 0;

Qy 1 ATGACATGACGGGAGGACGGCGCGGAGCGGCGGAGCGGCGGATGTTGGGCCCTACCGG 60
Db 387 ATGACATGACGGGAGGACGGCGCGGAGCGGCGGAGCGGCGGAGCGGCGGATGTTGGGCCCTACCGG 446

Qy 61 CTGGAGAGACGCTGGGCAAGGGGCGAGCAGGCTTGGTGAAGCTGGGGGTTCACCTGCGTC 120
Db 447 CTGGAGAGACGCTGGGCAAGGGGCGAGCAGGCTTGGTGAAGCTGGGGGTTCACCTGCGTC 506
Qy 121 ACCTGCGAGAGGTGGCCATCAAGATCGTCAACCGTGAAGAGCTCAGCAGATCGGTGCTG 180
Db 507 ACCTGCGAGAGGTGGCCATCAAGATCGTCAACCGTGAAGAGCTCAGCAGATCGGTGCTG 566
Qy 181 ATGAGGTGGAGGGGAGATCGGATCCTGAAGCTCATTTGAGCACCCCACTCTCTAAAG 240
Db 567 ATGAGGTGGAGGGGAGATCGGATCCTGAAGCTCATTTGAGCACCCCACTCTCTAAAG 626
Qy 241 CTGACAGAGTTTATGAAAAAATAATTTGTACTCTGCTGTAGAAACACGCTGTGAGGT 300
Db 627 CTGACAGAGTTTATGAAAAAATAATTTGTACTCTGCTGTAGAAACACGCTGTGAGGT 686
Qy 301 GGTGAGCTTTGAGCTACCTGTGAAAGAGGGAGGCTCAGCCTTAAGAGGCTCGGAAG 360
Db 687 GGTGAGCTTTGAGCTACCTGTGAAAGAGGGAGGCTCAGCCTTAAGAGGCTCGGAAG 746
Qy 361 TTCTTCCGCGCAGATCATCTCTGCGCTGGACTTCTGCCACAGCCACTCCATATGCCACAG 420
Db 747 TTCTTCCGCGCAGATCATCTCTGCGCTGGACTTCTGCCACAGCCACTCCATATGCCACAG 806
Qy 421 GATCTGAAAACCTGAAAACCTCTCTGTGAGAGAGAACAAACATCCGATCGCAGACTTT 480
Db 807 GATCTGAAAACCTGAAAACCTCTCTGTGAGAGAGAACAAACATCCGATCGCAGACTTT 866
Qy 481 GGATGCGGCTCTCGAGGTTCGCGAGCAGCTGTGAGACACAGCTGTGGTTCCTCCCGAC 540
Db 867 GGATGCGGCTCTCGAGGTTCGCGAGCAGCTGTGAGACACAGCTGTGGTTCCTCCCGAC 926
Qy 541 TAGCGCTGCCCGCAGGTGATTCGCGGGGAGAGTATGACGGCGGAGGCGGAGCTGTGG 600
Db 927 TAGCGCTGCCCGCAGGTGATTCGCGGGGAGAGTATGACGGCGGAGGCGGAGCTGTGG 986
Qy 601 AGCTGCGGCTCATCTCTGCTGCTGCTGGTGGGGGCTCTGCCCTTCGACGATGACAA 660
Db 987 AGCTGCGGCTCATCTCTGCTGCTGCTGGTGGGGGCTCTGCCCTTCGACGATGACAA 1046
Qy 661 TTGCGACAGCTGTGAGAGAGGTGAAGCGGGGGTGTTCACATGCGGCTTTATCCCG 720
Db 1047 TTGCGACAGCTGTGAGAGAGGTGAAGCGGGGGTGTTCACATGCGGCTTTATCCCG 1106
Qy 721 CCGGACTGCCCAGAGTCTGTACGGGGCATGATGAGGAGTGAAGCGGGGGTGTTCACATGCGGCTTTATCCCG 780
Db 1107 CCGGACTGCCCAGAGTCTGTACGGGGCATGATGAGGAGTGAAGCGGGGGTGTTCACATGCGGCTTTATCCCG 1166
Qy 781 CTAGAGCACATTCAGAAACACATATGTTATATAGGGGGCAAGAAATGAGCCGAAACAGAG 840
Db 1167 CTAGAGCACATTCAGAAACACATATGTTATATAGGGGGCAAGAAATGAGCCGAAACAGAG 1226
Qy 841 CAGGCCATTCCTGCAAGGTGAGATCCGCTCCGCGGAGCTGGAGGACATGACCC 900
Db 1227 CAGGCCATTCCTGCAAGGTGAGATCCGCTCCGCGGAGCTGGAGGACATGACCC 1286
Qy 901 GACGTCTGACACAGCTCACTCACTGCGGTGCTTCGAGACCGCAACAAAGCTGCTGAG 960
Db 1287 GACGTCTGACACAGCTCACTCACTGCGGTGCTTCGAGACCGCAACAAAGCTGCTGAG 1346
Qy 961 GACGTCTGCTCGAGGAGAGAACACAGGAGAGATGATTTCTCTCTCTCGGACCGG 1020
Db 1347 GACGTCTGCTCGAGGAGAGAACACAGGAGAGATGATTTCTCTCTCTCGGACCGG 1406
Qy 1021 AAAGAAAGTACCGGACCGAGGATGAGGACTGCCCCCCCCCGGACGAGATAGACCT 1080
Db 1407 AAAGAAAGTACCGGACCGAGGAGATGAGGACTGCCCCCCCCCGGACGAGATAGACCT 1466
Qy 1081 CCGCGAAGCGTGTGAGTCTCCCGATGCTGAAACCGGCAAGCGGCGGCGGAGACGC 1140
Db 1467 CCGCGAAGCGTGTGAGTCTCCCGATGCTGAAACCGGCAAGCGGCGGCGGAGACGC 1526
Qy 1141 AAATCCAAGAGGTGCTCAGCTGACGGACGGCGGTCTCCCGGTGCTGCGCGCGGGCC 1200

Db	1527	AAATTCATGGAGGTGCTCAGCGTACGAGACGCGGCTCCCGGTCCTGCGCGCGGCGC	1586
Qy	1201	ATTGAGATGGCCAGCAGCGGACAGAGTCTCGGTCCATCAGCGTGCTCTCTCAGGCTT	1260
Db	1587	ATTGAGATGGCCAGCAGCGGACAGAGTCTCGGTCCATCAGCGTGCTCTCTCAGGCTT	1646
Qy	1261	TCCACGAGCCCATCAGCAGCGGCGGCTGACCCCTTCCACCCCTTCCACGAGGCGAGTCCC	1320
Db	1647	TCCACGAGCCCATCAGCAGCGGCGGCTGACCCCTTCCACCCCTTCCACGAGGCGAGTCCC	1706
Qy	1321	CTCCCGACCCCGAGGAGACCTGTGTCCACGCGCAAGAGAGCCCGGCTGGCAGCGCCC	1380
Db	1707	CTCCCGACCCCGAGGAGACCTGTGTCCACGCGCAAGAGAGCCCGGCTGGCAGCGCCC	1766
Qy	1381	AACCCCGACCGCCCGTCCAGCGGCTCGGAGGGGTGCCCTGGAGGGCGCGCTCAAC	1440
Db	1767	AACCCCGACCGCCCGTCCAGCGGCTCGGAGGGGTGCCCTGGAGGGCGCGCTCAAC	1826
Qy	1441	TCCATCAAGAACAGCTTTCTGGGCTCACCCCGCTTCCACCGCGGAACTGCAAGTTCCG	1500
Db	1827	TCCATCAAGAACAGCTTTCTGGGCTCACCCCGCTTCCACCGCGGAACTGCAAGTTCCG	1886
Qy	1501	ACGCGGAGGAGATGTCCAACTGACACCGAGTGTCTCCAGAGCTGGCGAAGAGTCC	1560
Db	1887	ACGCGGAGGAGATGTCCAACTGACACCGAGTGTCTCCAGAGCTGGCGAAGAGTCC	1946
Qy	1561	TGGTTTGGAACTTCATCAGCGCTCGGAGGAGGAGGAGATCTTCTGGTGTATCAAGAG	1620
Db	1947	TGGTTTGGAACTTCATCAGCGCTCGGAGGAGGAGGAGATCTTCTGGTGTATCAAGAG	2006
Qy	1621	AAACCTCTGAGCTCCATCAAGGCTGACATCGTGCAAGCTTCTGTGATTCAGGCTTC	1680
Db	2007	AAACCTCTGAGCTCCATCAAGGCTGACATCGTGCAAGCTTCTGTGATTCAGGCTTC	2066
Qy	1681	AGCCACAGCTCATCTCCAAACAGCTTCGGGCGGAGTACAAAGCCAGCGGGGCGCA	1740
Db	2067	AGCCACAGCTCATCTCCAAACAGCTTCGGGCGGAGTACAAAGCCAGCGGGGCGCA	2126
Qy	1741	GCGGTGTTCCAGAGCGGCTCAAGTTCAGGTTCATATCACTTACCGAGGTTGGGGAG	1800
Db	2127	GCGGTGTTCCAGAGCGGCTCAAGTTCAGGTTCATATCACTTACCGAGGTTGGGGAG	2186
Qy	1801	GCGCAGAGGAGAACGGCATCTACTCCGTCACTTCACTCCCTGTCTCAGGCCCCAGCGT	1860
Db	2187	GCGCAGAGGAGAACGGCATCTACTCCGTCACTTCACTCCCTGTCTCAGGCCCCAGCGT	2246
Qy	1861	CGCTTCAAGAGGTTGGAGACCATCCAGGCCAGCTGTGTAGCACAACAGCCCGCT	1920
Db	2247	CGCTTCAAGAGGTTGGAGACCATCCAGGCCAGCTGTGTAGCACAACAGCCCGCT	2306
Qy	1921	GCGGCCAGCACTTGTTCAGACACCACTAACTGTATGGAATGATGAGCGGGCGGCTTCC	1980
Db	2307	GCGGCCAGCACTTGTTCAGACACCACTAACTGTATGGAATGATGAGCGGGCGGCTTCC	2366
Qy	1981	AAATGTGGAATTATCCCGAAAGTTAA	2007
Db	2367	AAATGTGGAATTATCCCGAAAGTTAA	2393

RESULT 4
AR232171
LOCUS AR232171 3364 bp DNA linear PAT 20-DEC-2002
DEFINITION Sequence 3 from patent US 6455292.
ACCESSION AR232171
VERSION AR232171.1 GI:27274062
KEYWORDS
SOURCE Unknown.
ORGANISM Unclassified.
REFERENCE 1 (bases 1 to 3364)
AUTHORS Shu,Y., Fan,W., Kovacs,K.F., Zidanic,M. and Jay,G.
TITLE Full-length serine protein kinase in brain and pancreas

JOURNAL FEATURES	Patent: US 6455292-A 3 24-SEP-2002;
source	Location/Qualifiers
	1..3364
BASE COUNT	639 a 1078 c 1089 g 578 t
ORIGIN	/organism="unknown"
	Query Match 99.3%; Score 1993; DB 6; Length 3364;
	Best Local Similarity 99.8%; Pred. No. 0;
	Matches 2007; Conservative 0; Mismatches 0; Indels 4; Gaps 1;
Qy	1 ATGACATTCAGCGGGAAGGACGCGCGCGCAGCAGCGCAGTATTTTGGGCCCTTACCGG 60
Db	229 ATGACATTCAGCGGGAAGGACGCGCGCGCAGCAGCGCAGTATTTTGGGCCCTTACCGG 288
Qy	61 CTGGAGAACAGCGTGGGCAAGGGGAGACAGAGTCTGCTGAAGCTGGGGTTCACTCGGTC 120
Db	289 CTGGAGAACAGCGTGGGCAAGGGGAGACAGAGTCTGCTGAAGCTGGGGTTCACTCGGTC 348
Qy	121 ACCTGCCAGAAGGTGGCCATCAAGATCGTCAACCGTGAGAAAGCTCAGCGAGTCEGTGCTG 180
Db	349 ACCTGCCAGAAGGTGGCCATCAAGATCGTCAACCGTGAGAAAGCTCAGCGAGTCEGTGCTG 408
Qy	181 ATGAAGTGGAGCGGAGATTCGCGATCTTGAAGCTCATTTAGACACCCCACTCTTAAAG 240
Db	409 ATGAAGTGGAGCGGAGATTCGCGATCTTGAAGCTCATTTAGACACCCCACTCTTAAAG 468
Qy	241 CTGCACAGCGTTTATGAACCAAAAATATTTT---GTACCTGGTGTAGAACACGCTGTC 296
Db	469 CTGCACAGCGTTTATGAACCAAAAATATTTTGTAGGTACCTGGTGTAGAACACGCTGTC 528
Qy	297 AGTGTGTGAGCTCTTCGACTTACCTGGTGAAAGAGGGAGGCTGACGCCCTAAAGGAGCTCG 356
Db	529 AGTGTGTGAGCTCTTCGACTTACCTGGTGAAAGAGGGAGGCTGACGCCCTAAAGGAGCTCG 588
Qy	357 GAAGTTCTTCCGCGAGATCATCTCTGCGCTGGACTTCTGCCACAGCCACTTCCATATGCCA 416
Db	589 GAAGTTCTTCCGCGAGATCATCTCTGCGCTGGACTTCTGCCACAGCCACTTCCATATGCCA 648
Qy	417 CAGGGATCTGAACCTGAAACCTCTCTGCGAGAGAGAGAACCAACATCCGATCCGAGA 476
Db	649 CAGGGATCTGAACCTGAAACCTCTCTGCGAGAGAGAGAACCAACATCCGATCCGAGA 708
Qy	477 CTTTGGCATGGGCTCCCTGCAGGTTGGCGACAGCTTGTGGAGACAGCTGTGGGTCCTCC 536
Db	709 CTTTGGCATGGGCTCCCTGCAGGTTGGCGACAGCTTGTGGAGACAGCTGTGGGTCCTCC 768
Qy	537 CCACTACGCTGCCCCGAGGTTGATCCGGGGGAGAAAGTATGACGGCCGGAAGCGGACGT 596
Db	769 CCACTACGCTGCCCCGAGGTTGATCCGGGGGAGAAAGTATGACGGCCGGAAGCGGACGT 828
Qy	597 GTGGAGCTCGGCGTCTCTCTGCTTCCCTTGTGGTGGGGGCTCTGCCCTTCGACGATGA 656
Db	829 GTGGAGCTCGGCGTCTCTCTGCTTCCCTTGTGGTGGGGGCTCTGCCCTTCGACGATGA 888
Qy	657 CAACCTTGCACAGCTCTCGAGAGAGGTGAAGCGGGGCTGTTCCACATGCCGCACTTTAT 716
Db	889 CAACCTTGCACAGCTCTCGAGAGAGGTGAAGCGGGGCTGTTCCACATGCCGCACTTTAT 948
Qy	717 CCGCCCCGACTGCGCAGAGTCTGCTACGGGGCATGATCGAGGTGGAACGCGCAGCGCGCT 776
Db	949 CCGCCCCGACTGCGCAGAGTCTGCTACGGGGCATGATCGAGGTGGAACGCGCAGCGCGCT 1008
Qy	777 CAGCTTAGAGCATTTCAGAAACACATATAGGGGGCAAGAAATGAGCCCAAC 836
Db	1009 CAGCTTAGAGCATTTCAGAAACACATATAGGGGGCAAGAAATGAGCCCAAC 1068
Qy	837 AGAGCAGCCCATCTCTCGAAGGTGAGATCCGCTCGCTGCCAGCTCGAGGACATCGA 896
Db	1069 AGAGCAGCCCATCTCTCGAAGGTGAGATCCGCTCGCTGCCAGCTCGAGGACATCGA 1128
Qy	897 CCGCAGCTGTGGACAGCATGCACTCACTGGGCTCTTCCGAGACCGCAACAGCTGCT 956

Db 1129 CCCCACGCTGTGGACAGCATGCACTCACTGGGCTGCTCCGAGACCGCAACAAGCTGCT 1188
Qy 957 GCAGGACCTGCTGTCGAGAGAGAGAACGAGAGAGATTTACTTCTCTCTCTGA 1016
Db 1189 GCAGGACCTGCTGTCGAGAGAGAGAACGAGAGAGATTTACTTCTCTCTCTGA 1248
Qy 1017 CCGGAAGAAAGTATCCCGAGCAGAGAGATGAGGACCTGCCCGCCCGGAAACAGATAGA 1076
Db 1249 CCGGAAGAAAGTATCCCGAGCAGAGAGATGAGGACCTGCCCGCCCGGAAACAGATAGA 1308
Qy 1077 CCTCTCCCGAAGCGTGTGACTCTCCCGATGCTGAACCGGACGCGCAAGCGCGGCAGA 1136
Db 1309 CCTCTCCCGAAGCGTGTGACTCTCCCGATGCTGAACCGGACGCGCAAGCGCGGCAGA 1368
Qy 1137 ACCCAATCCATGAGTGTGCTCAGCGTGACGAGCGCGCTCCCGTCTCTCGCGGG 1196
Db 1369 ACGCAATCCATGAGTGTGCTCAGCGTGACGAGCGCGCTCCCGTCTCTCGCGGG 1428
Qy 1197 GGCCATTGAGATGGCCAGACGCGCCAGAGGTCTCGTCCATCAGCGGTGCTCTCTCAGG 1256
Db 1429 GGCCATTGAGATGGCCAGACGCGCCAGAGGTCTCGTCCATCAGCGGTGCTCTCTCAGG 1488
Qy 1257 CTTTTCACAGCCACTCAGCAGCGCCCGGGTGACCCCTCACCCCTCAGCAAGGGGCGAG 1316
Db 1489 CTTTTCACAGCCACTCAGCAGCGCCCGGGTGACCCCTCACCCCTCAGCAAGGGGCGAG 1548
Qy 1317 TCCCTCCCAACCCCAAGGGGACACTGTCCACAGCCCAAGAGAGCCCGCTGCGCAC 1376
Db 1549 TCCCTCCCAACCCCAAGGGGACACTGTCCACAGCCCAAGAGAGCCCGCTGCGCAC 1608
Qy 1377 GCGCAACCCCAACCCCGCTCGAGCCCGAGCGTGGAGGGGTGCCCTGGAGGGCGCGGCT 1436
Db 1609 GCGCAACCCCAACCCCGCTCGAGCCCGAGCGTGGAGGGGTGCCCTGGAGGGCGCGGCT 1668
Qy 1437 CAATCTCATCAAGAAAGCTTTCTGGGCTCACCCGCTCCAGCGCGCGGAACTGCAAGT 1496
Db 1669 CAATCTCATCAAGAAAGCTTTCTGGGCTCACCCGCTCCAGCGCGGAACTGCAAGT 1728
Qy 1497 TCCGACGCGGAGAGATGTCCAACTTGACACAGAGTGTGTCGCCAGAGTGGCGAAGAA 1556
Db 1729 TCCGACGCGGAGAGATGTCCAACTTGACACAGAGTGTGTCGCCAGAGTGGCGAAGAA 1788
Qy 1557 GTCTGTGTTGGGAACCTTCATAGCTTGAGAGAGAGAGAGAGATCTTCTGTGTATCAA 1616
Db 1789 GTCTGTGTTGGGAACCTTCATAGCTTGAGAGAGAGAGAGAGATCTTCTGTGTATCAA 1848
Qy 1617 AGACAAACCTCTGAGCTGCATCAGGCTGACATCGTGCAAGCGCTTCTGTGTATCCAG 1676
Db 1849 AGACAAACCTCTGAGCTGCATCAGGCTGACATCGTGCAAGCGCTTCTGTGTATCCAG 1908
Qy 1677 TCTCAGGCACAGCGTCACTCCCAAGCAGCTTCCGGGCGGAGTACAAAGGCCACGGGGG 1736
Db 1909 TCTCAGGCACAGCGTCACTCCCAAGCAGCTTCCGGGCGGAGTACAAAGGCCACGGGGG 1968
Qy 1737 GCCAGCGTGTTCAGAAAGCGGTCAAGTTCAGGTTGATATACCTTACACGAGGGTGG 1796
Db 1969 GCCAGCGTGTTCAGAAAGCGGTCAAGTTCAGGTTGATATACCTTACACGAGGGTGG 2028
Qy 1797 GGAGGCGCAG 1856
Db 2029 GGAGGCGCAG 2088
Qy 1857 CCGTCTGCTTCAAGAGGGTGTGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1916
Db 2089 CCGTCTGCTTCAAGAGGGTGTGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2148
Qy 1917 GCCTGGGCGCCAGCACTTGTTCAGACACCACTTATGTAATGTAATGAGAGGGCGGCT 1976
Db 2149 GCCTGGGCGCCAGCACTTGTTCAGACACCACTTATGTAATGTAATGAGAGGGCGGCT 2208
Qy 1977 TTCCAAATGTGGAATTTATCCGGAAGTTAA 2007
Db 2209 TTCCAAATGTGGAATTTATCCGGAAGTTAA 2239

RESULT 5

AX327995
LOCUS AX327995 2025 bp DNA linear PAT 07-JAN-2002
DEFINITION Sequence 3 from Patent WO0181588.
ACCESSION AX327995
VERSION AX327995.1 GI:18098148
KEYWORDS Homo sapiens (human)
SOURCE Homo sapiens
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE

1 Meyers, R.
2246, protein kinase molecules and uses therefor
Patent: WO 0181588-A 3 01-NOV-2001;
Millennium Pharmaceuticals, Inc. (US)

FEATURES

1..2025
Location/Qualifiers
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
BASE COUNT 442 a 641 c 616 g 326 t
ORIGIN

Query Match 96.6%; Score 1939.6; DB 6; Length 2025;

Best Local Similarity 99.8%; Pred. No. 0;

Matches 1942; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 1 ATGACATCGACGGGGAAGGACGGCGCGCGAGCAGCGAGTATGTTGGGCGCTTACCGG 60
Db 1 ATGACATCGACGGGGAAGGACGGCGCGCGAGCAGCGAGTATGTTGGGCGCTTACCGG 60
Qy 61 CTGGAGAGACCTGGGCGAGGGGCGAGCAGAGTCTGGTGAAGCTGGGGTTTCACTGCGTC 120
Db 61 CTGGAGAGACCTGGGCGAGGGGCGAGCAGAGTCTGGTGAAGCTGGGGTTTCACTGCGTC 120
Qy 121 ACTGTCCAGAAAGTGGCCATCAAGATCGTCAACCGTGAGAGCTCAGCAGTCCGTTGCTG 180
Db 121 ACTGTCCAGAAAGTGGCCATCAAGATCGTCAACCGTGAGAGCTCAGCAGTCCGTTGCTG 180
Qy 181 ATGAAGTGGAGCGGGAGATTCGGATCTCTGAAGTCAATTGAGACCCCCACGTCCTAAAG 240
Db 181 ATGAAGTGGAGCGGGAGATTCGGATCTCTGAAGTCAATTGAGACCCCCACGTCCTAAAG 240
Qy 241 CTGCACGAGCTTTATGAAGAAACAAAAATATTTGTACCTGGTCTAGAACACGTTCAAGT 300
Db 241 CTGCACGAGCTTTATGAAGAAACAAAAATATTTGTACCTGGTCTAGAACACGTTCAAGT 300
Qy 301 GGTGAGCTCTTCGACTACCTCGTGAAGAGGGAGGCTGACGCTAAGGAGGCTCGGAAG 360
Db 301 GGTGAGCTCTTCGACTACCTCGTGAAGAGGGAGGCTGACGCTAAGGAGGCTCGGAAG 360
Qy 361 TTCTTCCGCGAGATCATCTCTGCGCTGGAATTCTGCGCACAGCCACTCCATATGCGCACAG 420
Db 361 TTCTTCCGCGAGATCATCTCTGCGCTGGAATTCTGCGCACAGCCACTCCATATGCGCACAG 420
Qy 421 GATCTGAACCTTGAACAACTCTCTGCTGGACGAGAGAAACAAATCCGATCGCAGACTTT 480
Db 421 GATCTGAACCTTGAACAACTCTCTGCTGGACGAGAGAAACAAATCCGATCGCAGACTTT 480
Qy 481 GGATCGGCGTCCCTGCGAGGTTGGCGACAGCCTGTTGGAGACGAGCTGTGGGTCCCCCCAC 540
Db 481 GGATCGGCGTCCCTGCGAGGTTGGCGACAGCCTGTTGGAGACGAGCTGTGGGTCCCCCCAC 540
Qy 541 TAGCCTGCGCGAGGTGATTCGGGGGAGAGATGATGACGGCGGAGGCGGAGCGAGTGTGG 600
Db 541 TAGCCTGCGCGAGGTGATTCGGGGGAGAGATGATGACGGCGGAGGCGGAGCGAGTGTGG 600
Qy 601 AGCTGCGGCGTCACTCTGCTTCCCTTGGTGGGGGCTCTGCCCTTCGACGATGACAAAC 660
Db 601 AGCTGCGGCGTCACTCTGCTTCCCTTGGTGGGGGCTCTGCCCTTCGACGATGACAAAC 660

```
QY 661 TTGCGACAGCTGCTGGAGAGGTGAAGCGGGCGTGTCCACATGCCCGCACTTTATCCCG 720
DB 661 TTGCGACAGCTGCTGGAGAGGTGAAGCGGGCGTGTCCACATGCCCGCACTTTATCCCG 720
QY 721 CCGGACTGCGAGAGTCTGTACGGGGGCAATGATGAGAGTGGAGCGCGCAGCGCTCAAG 780
DB 721 CCGGACTGCGAGAGTCTGTACGGGGGCAATGAGAGTGGAGCGCGCAGCGCTCAAG 780
QY 781 CTAGAGCACATTTCAAGAACACATATGTTATATAGGGGGCAAGATGAGCGGACCAAG 840
DB 781 CTAGAGCACATTTCAAGAACACATATGTTATATAGGGGGCAAGATGAGCGGACCAAG 840
QY 841 CAGCCCATTTCTCGCAAGGTGCAGATCGCTCGCTGCCAGCCTGAGGAGCATCGACCCC 900
DB 841 CAGCCCATTTCTCGCAAGGTGCAGATCGCTCGCTGCCAGCCTGAGGAGCATCGACCCC 900
QY 901 GACGTGCTGGACAGATGCACTACTGGGCTGCTTCCGAGACCGCAACAAGTGTGAG 960
DB 901 GACGTGCTGGACAGATGCACTACTGGGCTGCTTCCGAGACCGCAACAAGTGTGAG 960
QY 961 GACCTGCTGTCGAGAGGAGAACACAGAGAGATGATTTACTTCTCTCTCTGAGACCG 1020
DB 961 GACCTGCTGTCGAGAGGAGAACACAGAGAGATGATTTACTTCTCTCTCTGAGACCG 1020
QY 1021 AAAGAAAGGTATCCGAGGCGAGAGATGAGGACCTGCCCCCGGAAACGAGATAGACCT 1080
DB 1021 AAAGAAAGGTATCCGAGGCGAGAGATGAGGACCTGCCCCCGGAAACGAGATAGACCT 1080
QY 1081 CCGCGAAGCGTGTGGAATCTCCCGATGCTGAACCGGCAAGCGCGGCGGCAAGACGC 1140
DB 1081 CCGCGAAGCGTGTGGAATCTCCCGATGCTGAACCGGCAAGCGCGGCGGCGGCAAGACGC 1140
QY 1141 AAATCCATGAGGTGCTCAGCGTGACGACGCGGCTCCCGGTGCTGCGCGGCGGCG 1200
DB 1141 AAATCCATGAGGTGCTCAGCGTGACGACGCGGCTCCCGGTGCTGCGCGGCGGCG 1200
QY 1201 ATTGAGATGGCCAGCACGGCCAGAGGTCTCGGTCCATCAGCGGTGCTCTCTCAGGCGCTT 1260
DB 1201 ATTGAGATGGCCAGCACGGCCAGAGGTCTCGGTCCATCAGCGGTGCTCTCTCAGGCGCTT 1260
QY 1261 TCCACAGGCGCACTCAGACGCGCCCGGTGACCTCTACCCCTCAACCAAGGGGCAAGTCCC 1320
DB 1261 TCCACAGGCGCACTCAGACGCGCCCGGTGACCTCTACCCCTCAACCAAGGGGCAAGTCCC 1320
QY 1321 CTCGCCACCCCAAGSGGACACTCTGCACAGCCCAAGAGGAGCGCGGTGGCGACGCCC 1380
DB 1321 CTCGCCACCCCAAGSGGACACTCTGCACAGCCCAAGAGGAGCGCGGTGGCGACGCCC 1380
QY 1381 AAACCCAGCGCCCGCTCCAGCGCCAGCGTGGAGGGGTGCCCTGGAGGGCGCGGCTCAAC 1440
DB 1381 AAACCCAGCGCCCGCTCCAGCGCCAGCGTGGAGGGGTGCCCTGGAGGGCGCGGCTCAAC 1440
QY 1441 TCATCAGAAACAGCTTTCTGGGCTCACCCCGCTTCCACCGCGGAAACTGCAAGTTCCG 1500
DB 1441 TCATCAGAAACAGCTTTCTGGGCTCACCCCGCTTCCACCGCGGAAACTGCAAGTTCCG 1500
QY 1501 AGCGGAGGAGATGTCACACTGACACAGAGTCTCCCGAGAGTCCCGAGAGTCCCGAGTCC 1560
DB 1501 AGCGGAGGAGATGTCACACTGACACAGAGTCTCCCGAGAGTCCCGAGAGTCCCGAGTCC 1560
QY 1561 TGGTTTGGGAACTTCATCAGCTGGAGAGGAGGAGAGATCTTCTGTTGATCAAGAGAC 1620
DB 1561 TGGTTTGGGAACTTCATCAGCTGGAGAGGAGGAGAGATCTTCTGTTGATCAAGAGAC 1620
QY 1621 AAACCTCTGAGTCTCATCAAGGCTGACATGCTGCAAGCTTCTCTCTGATCCAGTCTC 1680
DB 1621 AAACCTCTGAGTCTCATCAAGGCTGACATGCTGCAAGCTTCTCTCTGATCCAGTCTC 1680
QY 1681 AGCCACAGCTCATCTCCCAACAGCTTCCGGGCGGAGTCAAGGCGCACCGGGGGGCA 1740
DB 1681 AGCCACAGCTCATCTCCCAACAGCTTCCGGGCGGAGTCAAGGCGCACCGGGGGGCA 1740
```

```
QY 1741 GCCGTGTTCCAGAACGCGGTCAAGTTCACGTTGATATACCTACACGAGGGTGGGGAG 1800
DB 1741 GCCGTGTTCCAGAACGCGGTCAAGTTCACGTTGATATACCTACACGAGGGTGGGGAG 1800
QY 1801 GCGGAGAGGAGAACGCGATCTACTTCGCTCACTTCACTTCACTTCACTTCACTTCACTT 1860
DB 1801 GCGGAGAGGAGAACGCGATCTACTTCGCTCACTTCACTTCACTTCACTTCACTTCACTT 1860
QY 1861 CGCTTCAAGAGGGTGGTGGAGACCATCCAGCCCGAGCTGCTGAGCACACAGACCCGCT 1920
DB 1861 CGCTTCAAGAGGGTGGTGGAGACCATCCAGCCCGAGCTGCTGAGCACACAGACCCGCT 1920
QY 1921 GCGGCCAGCAGCTTGTGAGACACAC 1946
DB 1921 GCGGCCAGCAGCTTGTGAGACACAC 1946

RESULT 6
AX327993 LOCUS 2219 bp DNA linear PAT 07-JAN-2002
DEFINITION Sequence 1 from Patent WO0181588.
ACCESSION AX327993
VERSION AX327993.1 GI:18098146
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1
AUTHORS Meyers,R.
TITLES 2246, protein kinase molecules and uses therefor
JOURNAL Patent: WO 0181588-A 1 01-NOV-2001;
Millennium Pharmaceutical, Inc. (US)
FEATURES
source
1. 2219
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
53..2077
/note="unnamed protein product"
/codon_start=1
/protein_id="CAD20436.1"
/db_xref="GI:18098147"
/translation="MTSTGDKGAQAQYVGPYRLEKTLGKGTGLVKLGVCVTCK
VAIKVREKLSSEVLKMKVERETAILKLEHVLKLDVYENKXVLYLVLEHVSGB
LFDYLVKGLRTPKEARKPFRQIISALDFCHSHSICHRDLKPNLLDKNNRIADP
GMASLQVDSILLETSCSPHYACPEVIRGEKYDGRKADVMSGVLLFALLVGLPFD
DNLRQLLEKRYGVFTHMPHFIPDQSLRGMSEVDAARLTLEHQQHWTYGGKNE
PEPEQTPRQVQIRSLPDLIDPDLDSMHSIGCFDRNKLQDLLESEENQEKMY
FLLLDRKERYPSQEDLPNRNIDPPRKYVDSPLNLRHGRKRPERSKMSVLSVTDG
SPVPARRAISMACHGORSRSISGASGLSTSPSSPRVTPHSPRGSPLPTPKGTPVH
TPKESPATNPTPTPSPSPSVGVWRARLNSIKNSFLGSPFRHRLQVPTPEMSNL
TPSSPELAKGWGFWGIFISLEKEEQIFVVDKPLSKADI VHAFLUSIPLSHSVLS
QTSFRAFYKATGGAFFQKVFQVDITYTEGGEAQKNGIYSVTFLLSGPSRRPKR
VVTETQQLLSLTHDPPAAQHLSEPPPPAPGLSWAGLKGQKQVATSYESSL"
BASE COUNT 476 a 716 c 683 g 344 t
ORIGIN
Query Match 96.6%; Score 1939.6; DB 6; Length 2219;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 1942; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
QY 1 ATGACATCGACGGGAGGACGGCGCGCACACGCGCAGTATGTGGGCCCTACCGG 60
DB 53 ATGACATCGACGGGAGGAGCGCGCGCGCAGCAGCGCAGTATGTGGGCCCTACCGG 112
QY 61 CTGAGAGAGCGCTGGGCAAGGGGCGACAGAGTCTGTTGAAGCTGGGGGTTCACTGGCTC 120
DB 113 CTGAGAGAGCGCTGGGCAAGGGGCGACAGAGTCTGTTGAAGCTGGGGGTTCACTGGCTC 172
QY 121 ACTTGCCAGAGGTGGCCATCAAGATCGTCAACCGTGAGAGGCTCAGCGAGTGGTCTG 180
DB 173 ACTTGCCAGAGGTGGCCATCAAGATCGTCAACCGTGAGAGGCTCAGCGAGTGGTCTG 232
```


Query Match 91.0%; Score 1827; DB 6; Length 1827;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1827; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	181	ATGAAGTGGAGCGGAGATCCGATCTCTGAAGCTCATTGAGCACCCGCCAGTCTTAAG	240
DB	1	ATGAAGTGGAGCGGAGATCCGATCTCTGAAGCTCATTGAGCACCCGCCAGTCTTAAG	60
QY	241	CTGACGACGCTTTATGAACAAACAAATATTTGTACTGTGTAGAACACGTTGACGGT	300
DB	61	CTGACGACGCTTTATGAACAAACAAATATTTGTACTGTGTAGAACACGTTGACGGT	120
QY	301	GGTGAAGTCTTGAAGTCTGTAAGTGAAGGAGGCTGACGCTTAAGAGAGCTCGAAG	360
DB	121	GGTGAAGTCTTGAAGTCTGTAAGTGAAGGAGGCTGACGCTTAAGAGAGCTCGAAG	180
QY	361	TTCTTTCCGGCAGATCATCTCTGCGCTGACATTTGTCACACGCACTCCATATGCCACAGG	420
DB	181	TTCTTTCCGGCAGATCATCTCTGCGCTGACATTTGTCACACGCACTCCATATGCCACAGG	240
QY	421	GATCTGAACCTGAAACCTCTCTGTCGACGAGAGAAACAAACATCCGATCGCAGACTTT	480
DB	241	GATCTGAACCTGAAACCTCTCTGTCGACGAGAGAAACAAACATCCGATCGCAGACTTT	300
QY	481	GGCATGGCGTCCCTGACAGGTTGGACAGCTGTTGGAGACAGCTGTGGTCCGCCAC	540
DB	301	GGCATGGCGTCCCTGACAGGTTGGACAGCTGTTGGAGACAGCTGTGGTCCGCCAC	360
QY	541	TACGCTCCGCCGAGTGCATCCGGGGGAGAGATGACGGCCGGAGGCGGAGCTGTGG	600
DB	361	TACGCTCCGCCGAGTGCATCCGGGGGAGAGATGACGGCCGGAGGCGGAGCTGTGG	420
QY	601	AGCTGCGCGCTCATCTCTGTCGCTTGTGTGGGGGCTCTGCGCTTCGACGATGACAAAC	660
DB	421	AGCTGCGCGCTCATCTCTGTCGCTTGTGTGGGGGCTCTGCGCTTCGACGATGACAAAC	480
QY	661	TTGGACAGCTGCTGGAGAGTGAAGCGGGCGTGTTCACATGCCGCACTTATCCCG	720
DB	481	TTGGACAGCTGCTGGAGAGTGAAGCGGGCGTGTTCACATGCCGCACTTATCCCG	540
QY	721	CCGACTGCGAGAGTGTCTACGGGGCATGATCGAGGTGAGCGCGGCGGCTCAG	780
DB	541	CCGACTGCGAGAGTGTCTACGGGGCATGATCGAGGTGAGCGCGGCGGCTCAG	600
QY	781	CTAGAGCATTTCAGAAACACATATGTTATAGGGGGCAAGATGAGCCCGAACACAG	840
DB	601	CTAGAGCATTTCAGAAACACATATGTTATAGGGGGCAAGATGAGCCCGAACACAG	660
QY	841	CAGCCCATTCCTGCAAGGTGAGATCCGCTGCTGCTGCTGCTGAGGAGATCGACCCC	900
DB	661	CAGCCCATTCCTGCAAGGTGAGATCCGCTGCTGCTGCTGCTGAGGAGATCGACCCC	720
QY	901	GACGTGCTGGACAGCATGCACTCTGCTGCTGCTTCCGAGACCGCAACAGCTGTGAG	960
DB	721	GACGTGCTGGACAGCATGCACTCTGCTGCTGCTTCCGAGACCGCAACAGCTGTGAG	780
QY	961	GACCTGCTGCTCCGAGGAGAACACAGGAGATGATTTACTTCTCTCTCTGACCGG	1020
DB	781	GACCTGCTGCTCCGAGGAGAACACAGGAGATGATTTACTTCTCTCTCTGACCGG	840
QY	1021	AAAGAAAGTATCCGAGCAGGAGATGAGACCTGCCCGCCCGGAAACGAGATAGACCT	1080
DB	841	AAAGAAAGTATCCGAGCAGGAGATGAGACCTGCCCGCCCGGAAACGAGATAGACCT	900
QY	1081	CCCCGGAAGCTGTGACATCCCGGATGCTGAACCGGACGCGCAAGCGCGCAGAACGC	1140
DB	901	CCCCGGAAGCTGTGACATCCCGGATGCTGAACCGGACGCGCAAGCGCGCAGAACGC	960
QY	1141	AAATCCATGAGGTTGCTCAGCGTGAAGCGCGGCTCCCGGTTGCTGCGCGGGGCC	1200
DB	961	AAATCCATGAGGTTGCTCAGCGTGAAGCGCGGCTCCCGGTTGCTGCGCGGGGCC	1020
QY	1201	ATTGAGTGGCCACGACGGCCAGAGGTTCTCGGTCCATCAGCGGTGCTCTCAGGCGCTT	1260

DB	1021	ATTGAGTGGCCACGACGCGCCAGAGGTTCTCGGTCCATCAGCGGTGCTCTCAGGCGCTT	1080
QY	1261	TCCACACAGCCCTCAGCAGCGCCCGGGTGACCCCTCACCCTCACCAGGGGAGTCCC	1320
DB	1081	TCCACACAGCCCTCAGCAGCGCCCGGGTGACCCCTCACCCTCACCAGGGGAGTCCC	1140
QY	1321	CTCCCAACCCCAAGGGGACACCTGTCCACACGCCAAAGAGAGCGGCTGGACGCCC	1380
DB	1141	CTCCCAACCCCAAGGGGACACCTGTCCACACGCCAAAGAGAGCGGCTGGACGCCC	1200
QY	1381	AACCCACGCGCCCTCCAGCCCGGCTGAGAGGGTGCCCTGAGAGGGCGGCTCAAC	1440
DB	1201	AACCCACGCGCCCTCCAGCCCGGCTGAGAGGGTGCCCTGAGAGGGCGGCTCAAC	1260
QY	1441	TCCATCAAGAAACAGCTTTCTGGGCTCACCCCGCTTCCACCGCGGAAACTGCAAGTTCCG	1500
DB	1261	TCCATCAAGAAACAGCTTTCTGGGCTCACCCCGCTTCCACCGCGGAAACTGCAAGTTCCG	1320
QY	1501	ACGCCGAGGAGATGTCCAACTGACACCAAGAGTCTGTCGCCAGAGCTGGCGAAGTCC	1560
DB	1321	ACGCCGAGGAGATGTCCAACTGACACCAAGAGTCTGTCGCCAGAGCTGGCGAAGTCC	1380
QY	1561	TGTTTGGGAACTTCAATCAGCCTGGAGAGGAGGAGATCTTCTGTTGATCAAGAGAC	1620
DB	1381	TGTTTGGGAACTTCAATCAGCCTGGAGAGGAGGAGATCTTCTGTTGATCAAGAGAC	1440
QY	1621	AAACCTCTGAGCTCCATCAAGGCTGACATCGTGACGCTTCTGTCGATTTCCAGTCTC	1680
DB	1441	AAACCTCTGAGCTCCATCAAGGCTGACATCGTGACGCTTCTGTCGATTTCCAGTCTC	1500
QY	1681	AGCCACAGCGTCTATCTCCAAACGAGCTTCCGGGCGGAGTACAAGGCCACGGGGGGCCA	1740
DB	1501	AGCCACAGCGTCTATCTCCAAACGAGCTTCCGGGCGGAGTACAAGGCCACGGGGGGCCA	1560
QY	1741	GCCTGTTCCAGAGCGGTCAGTTCCAGGTTGATATCACTTACACGAGGGTGGGGAG	1800
DB	1561	GCCTGTTCCAGAGCGGTCAGTTCCAGGTTGATATCACTTACACGAGGGTGGGGAG	1620
QY	1801	GCGCAGAGGAGAACGGCATCTACTCCGTCACCTTCACTTCACTTCACTTCACTTCACTT	1860
DB	1621	GCGCAGAGGAGAACGGCATCTACTCCGTCACCTTCACTTCACTTCACTTCACTTCACTT	1680
QY	1861	CGTTTCAAGAGGTTGGAGACATCCAGCGCCAGCTGTGAGCACACGACCCGCT	1920
DB	1681	CGTTTCAAGAGGTTGGAGACATCCAGCGCCAGCTGTGAGCACACGACCCGCT	1740
QY	1921	GCGGCCAGCACTTGTGAGACACCACTAATCTGATGGAATGATGAGCGGGCGCTTCC	1980
DB	1741	GCGGCCAGCACTTGTGAGACACCACTAATCTGATGGAATGATGAGCGGGCGCTTCC	1800
QY	1981	AAATGTGAATTTATCCCGAAAGTTAA	2007
DB	1801	AAATGTGAATTTATCCCGAAAGTTAA	1827

RESULT 8

LOCUS	HSA6701
DEFINITION	Homo sapiens mRNA for putative serine/threonine protein kinase,
ACCESSION	AJ006701.1
VERSION	GI:3217027
KEYWORDS	putative; serine/threonine protein kinase.
SOURCE	Homo sapiens (human)
ORGANISM	Homo sapiens
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS	Mammalia; Euthera; Primates; Catarrhini; Hominidae; Homo.
TITLE	Stanchi, F., Bertocco, E., Toppo, S., Dioguardi, R., Simonati, B., Cannata, N., Zimbello, R., Lanfranchi, G. and Valle, G. Characterization of 16 novel human genes showing high similarity to yeast sequences

JOURNAL	Yeast 18 (1), 69-80 (2001)	QY	692	GGGTGTTCCACATCGCGACATTTATCCGCGCCGACTCTCCAGAGTCTGCTACCGGGCATGA	751
MEDLINE	21064499	DB	481	GGGTGTTCCACATCGCGACATTTATCCGCGCCGACTCTCCAGAGTCTGCTACCGGGCATGA	540
PUBMED	11124703				
REFERENCE	2 (bases 1 to 1956)	QY	752	TCAGAGTGGACCGCCGACCGCCCTCAGCTAGACACATTCCAGAAACATATGATATA	811
AUTHORS	Stanchi, F.	DB	541	GGAGGTGGACCGCCGACCGCCCTCAGCTAGACACATTCCAGAAACATATGATATA	600
TITLE	Direct Submission	QY	812	TAGGGGGGGAAGATGAGCCGGAACGAGACGAGCCCATCTCCGAGGTGCGAGATCCGCT	871
JOURNAL	Universita' di Padova, Via G. Colombo 3, 35121 Padova, 35121, ITALY	DB	601	TAGGGGGGGAAGATGAGCCGGAACGAGACGAGCCCATCTCCGAGGTGCGAGATCCGCT	660
FEATURES	Location/Qualifiers	QY	872	CGCTGCCAGAGCTGGAGGACATCGACCCCGAGCTGCTGAGACAGCATCACTCACTGGGCT	931
source	1. .1956	DB	661	CGCTGCCAGAGCTGGAGGACATCGACCCCGAGCTGCTGAGACAGCATCACTCACTGGGCT	720
	/organism="Homo sapiens"	QY	932	GTTTCGAGACCGCAACAAAGCTGCTGAGGACCTGCTGTCGAGGAGGAGAACCCAGGAGA	991
	/mol_type="mRNA"	DB	721	GCTTCGAGACCGCAACAAAGCTGCTGAGGACCTGCTGTCGAGGAGGAGAACCCAGGAGA	780
	/db_xref="taxon:9606"	QY	992	AGATGATTTTACTTCTCTCTCTGAGCCGGAAGAAAGGTACCCGAGCCAGGAGGATGAGG	1051
	/sex="female"	DB	781	AGATGATTTTACTTCTCTCTCTGAGCCGGAAGAAAGGTACCCGAGCCAGGAGGATGAGG	840
	/tissue_type="whole brain"	QY	1052	ACCTGCCCCCGGAAACGAGATAGACCTCTCCCGGAAGCGTGTGGACTCCCGAGTCTGA	1111
	/clone_lib="Soares infant brain INIB"	DB	841	ACCTGCCCCCGGAAACGAGATAGACCTCTCCCGGAAGCGTGTGGACTCCCGAGTCTGA	900
	/dev_stage="73 days post natal"	QY	1112	ACCGGACCGGCAAGCGGCGCGCAGAACGCAAAATCCATGAGAGTGTCTGACGTGACGGAG	1171
	<1. .1814	DB	901	ACCGGACCGGCAAGCGGCGCGCAGAACGCAAAATCCATGAGAGTGTCTGACGTGACGGAG	960
	/codon_start=3	QY	1172	GGGGTCTCCCGGCTGCTGCGCGCGCGGCGCATTTGAGATGGCCCGCAGACGCGGAGGTCTC	1231
	/product="putative serine/threonine protein kinase"	DB	961	GGGGTCTCCCGGCTGCTGCGCGCGCGGCGCATTTGAGATGGCCCGCAGACGCGGAGGTCTC	1020
	/protein_id="CAA07196.1"	QY	1232	GGTCCATCAGCGGTGCTCTCTCAGGCTTTTCCACAGCCCACTCAGACGCCCCCGGGTGA	1291
	/db_xref="GI:3217028"	DB	1021	GGTCCATCAGCGGTGCTCTCTCAGGCTTTTCCACAGCCCACTCAGACGCCCCCGGGTGA	1080
	/translation="LIEHPVLKLDVYENKYLILVLEHVSQGELEFDLVKKGRITP	QY	1292	CCCTCACCCTTCAACAGGGGCGAGTCCCTCCACACCCCGCCAGGGGACACCTGTCCACA	1351
	KEARKFRQIISALDFEHSIICHRDLKPNLLLDKNNIRIADFGWASLYQVDSLLLE	DB	1081	CCCTCACCCTTCAACAGGGGCGAGTCCCTCCACACCCCGCCAGGGGACACCTGTCCACA	1140
	TSCGSPHYACPEVIRGKDYGRKADVMSGVILFALLVAGALPFDNDLROLLEKVRG	QY	1352	CGCCAAAGGAGAGCGCGGTGCGACGCCCAACGCCGCGCCGCTCCAGCCCGCGGTCG	1411
	VFHMPHPIPDQCSSLRGMSVDAAERLTLEHIOKHIWYIGKNEPEPQPIPRKQVI	DB	1141	CGCCAAAGGAGAGCGCGGTGCGACGCCCAACGCCGCGCCGCTCCAGCCCGCGGTCG	1200
	RLSPLEIDIPDQVDSLWSHSLGCFADRNLQLDLLEENQRMVIFILLDKRIYFSQ	QY	1412	GAGGGTGGCTTGGAGGGCGCGGCTCAACTCCATCAAGAAACAGCTTTCTGGGCTCACCCC	1471
	HEDELPRNIDIPPRKRVDFMLNRHGRKRPERSMEVLSVTGGSPVFPARRAIEWAQ	DB	1201	GAGGGTGGCTTGGAGGGCGCGGCTCAACTCCATCAAGAAACAGCTTTCTGGGCTCACCCC	1260
	GGQSRISISGASGLSTSPSSPRVTPHPSPRGSPLTPKGTPTVTPKSPAGTNPPT	QY	1472	GCTTCCACCGCGGAAACTGCAAGTTCGACCGCGGAGAGATGTCCAACTGACACCAAG	1531
	PPSPSVGGVPMRLARLNSIKNSFLGSRPHRRKLOVTPPEMSNLTPSPSPPELAKSW	DB	1261	GCTTCCACCGCGGAAACTGCAAGTTCGACCGCGGAGAGATGTCCAACTGACACCAAG	1320
	FGNFIKLEKEQIPVVIKDKPLSLKADIVHAFSLPSLSHSVLSOTSPRAEYKATGG	QY	1532	AGTGTGTCCTCCAGAGCTGGGAAAGAGTCTGTTGGGAACTTTCATCAGCTGGGAGAGG	1591
	PAVQKPVKQVDITYTEGEAOKENGIYSVTLLSGSPRRPRKRVETIQAQLLSTH	DB	1321	AGTGTGTCCTCCAGAGCTGGGAAAGAGTCTGTTGGGAACTTTCATCAGCTGGGAGAGG	1380
	DPFAQLHSEFPFPAPGLSGWGLGKQVATSYESSL"	QY	1592	AGGAGCAGATCTTCTGTGCTCATCAAGACAAACCTCTGAGCTCCCATCAAGGCTGACATCG	1651
	427 a 640 c 579 g 310 t	DB	1381	AGGAGCAGATCTTCTGTGCTCATCAAGACAAACCTCTGAGCTCCATCAAGGCTGACATCG	1440
BASE COUNT		QY	1652	TGACGCGCTTCTGTGCTCATCCAGTCTCAGCCACAGCGTCTATCTCCCAACGAGGTTCC	1711
ORIGIN		DB	1441	TGACGCGCTTCTGTGCTCATCCAGTCTCAGCCACAGCGTCTATCTCCCAACGAGGTTCC	1500
		QY	1712	GGGCGCAGGTACAAAGGCGACGGGGGCGCAGCGTGTTCAGAAAGCGGTCAGGTTCCAGG	1771
		DB	1501	GGGCGCAGGTACAAAGGCGACGGGGGCGCAGCGTGTTCAGAAAGCGGTCAGGTTCCAGG	1560
		QY	1772	TTGATATACCTACACGAGGGGTGGGGGAGCGCGAGAAAGGAGAACCGGCATCTACTCCGTCA	1831

QY	1659	CTTCTCTGTCGATTCCAGTCTCAGCCACACAGCGTCATCTCCCAAACGAGCTTCCTCGGCGCCGA	1718
Db	1168	CTTCTCTGTCGATTCCAGTCTCAGCCACACAGCGTCATCTCCCAAACGAGCTTCCTCGGCGCCGA	1227
QY	1719	GTACAAGGCCACGGGGGGGCCAGCGCTGTTCCAGAGAGCGGTCAAGTTCAGGTTTCATAT	1778
Db	1228	GTACAAGGCCACGGGGGGGCCAGCGCTGTTCCAGAGAGCGGTCAAGTTCAGGTTTCATAT	1287
QY	1779	CACCTACACGGAGGGTGGGGAGCGCAGAGAGAGAACCGGCATCTACTCCGTCACTTCAC	1838
Db	1288	CACCTACACGGAGGGTGGGGAGCGCAGAGAGAGAACCGGCATCTACTCCGTCACTTCAC	1347
QY	1839	CTTGCTCTCAGGCCCCCAGCGCTCGCTTCAAGAGGGTGGTGGAGACCATTCAGGCCCAGCT	1898
Db	1348	CTTGCTCTCAGGCCCCCAGCGCTCGCTTCAAGAGGGTGGTGGAGACCATTCAGGCCCAGCT	1407
QY	1899	GCTGAGCACACAGCCCGCTCGGGCCAGCAGCTTCTCAGACACCCAC	1946
Db	1408	GCTGAGCACACAGCCCGCTCGGGCCAGCAGCTTCTCAGACACCCAC	1455
RESULT 10			
AF479827			
LOCUS	AF479827	3109 bp	linear
DEFINITION	Homo sapiens protein kinase-like protein mRNA		PRI 12-MAR-2002
ACCESSION	AF479827		
VERSION	AF479827.1	GI:19401873	
KEYWORDS			
SOURCE	Homo sapiens (human)		
ORGANISM	Homo sapiens		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.		
AUTHORS	She, X.Y., Yu, L. and Guo, J.H.		
TITLE	Direct Submission		
JOURNAL	Submitted (02-FEB-2002) School of Life Sciences, Laboratory of Human Genes Research, Institute of Genetics, Fudan University, 220 Handan Road, Shanghai 200433, P. R. China		
FEATURES	Location/Qualifiers		
source	1..3109		
	/organism="Homo sapiens"		
	/mol_type="mRNA"		
	/db_xref="taxon:9606"		
	/chromosome="19"		
	278..2614		
CDS	/codon_start=1		
	/product="protein kinase-like protein"		
	/protein_id="AAU87698.1"		
	/db_xref="GI:19401874"		
	/translation="MSSGAKEGGGSPAYLPHPHPPHQAQVGVPRLEKTLGKQ TGLVKLVGHCTGGQKVAIKVNRKSESVMKVERIEAILKLTIEHPVLKLDHVDYPR KYLYLVLEHVSQGLFDYLVKGRLLTPKEARFQIVSALDFCHSYSTCHRLKPRF NILLDEKNTRIADFVMSLIQVDSLLETSCGSPHYACPEVIGKEKYDGRADNWSGQ VQIFALLVGLPFDDDLRLLEKVRKGVFHMHPFITPDCOSLLRGMIEVEPEKRLRS EHAFLYVLLGGKHEPDPCLEPAPGRVAMRSLPSNGELDPDLVLESMAISGLCFDRDR RPERKSENEVKSIYLLLDLRKERYPCSDQDLPRLNVDPPKRVDSFMLSRRKPS RYRERKSMELVSIITDAGGSGSPVTRRALEMAHQSRVSVSAGTSLGSSPLSPGSR PVSFSPGAGDEARGGGSPSTKTQTLPSRGGGAGGQPPPSARSPTPLGPPGSGS PVRSGPPLHSPPLASPTPTGTTTPPSPGGVGGAARSLNIRNSFLGSPRRF HRSMQVTPAEWSLTPSPESPELAKSWFNGFISLDEKEQIFLVLDKDLPSSTKADI VQAFGLSTPSLHVSLSOTSFRAEYKASGSPVQKPVRFQVDTSSSEGGPSPRRDGS GGGGISTVFTLISGGRRRVVETIQALLSTHDQPSVQALADENKGAQTRPAGAP PRSIQPPGPRDPDPELSSPRRGPDKLLATNGTLP"		
BASE COUNT	611 a 1033 c 957 g 508 t		
ORIGIN			
Query Match	48.1%	Score 965;	DB 9; Length 3109;
Best Local Similarity	70.0%;	Pred. No. 2.e-162;	
Matches 1484;	Conservative 0;	Mismatches 440;	Indels 195; Gaps 6;
QY	29	CGCAGCACGCGACGATGTTGGCGCCCTACCGGCTGGAGAACGACGCTGGCGAAGGGCGAGA	98

351	CCAGCAGCCCCAAATATGTGGCCCTTATCGGCTGGAGAGACGCTGGGCAAAAGCACAGA	410
89	CAGGTCTGTGTGAAGCTGGGGGTTCATCTCGTCACTGCGCAGAAAGTGGCCATCAAGATCG	148
411	CAGGGCTGGTTAAACTCTGGGGTCCACTGTCATCACGGTGAGAAAGTTCGCCATCAAGATCG	470
149	TCAAACGGTGAAGAGCTCAGCGAGTCTGGTCTGTGATGAAGTGGAGCGGGAGATCGCGATCC	208
471	TCGAAACGGGAGAGAGTGTGGAGTGGGTGCTGATGAAGGTGGAGCGGGAGATCGCGATCC	530
209	TGAAGCTCAATTGAGCAGCCCCACGCTCTTAAAGCTGCACGACGTTTATGAAAAAATAAAT	268
531	TGAGGCTCATCGAACACCCACATGCTCTCAAGTCCAAGCTCTACGAGAACAGAAAT	590
269	ATTGTGACTGTGTGTAGAAACAGTGTCAAGTGTGAGCTCTTCGACTACCTGGTGAAGA	328
591	ATTGTGACTGTGTGTGAGCAGCTCTCGGGGGTGAGCTATTGCACTACCTGGTAAAGA	650
329	AGGGAGGCTGACGCTTAAGGAGGCTCGGAAGTTCTTCGGGAGATCATCTTCGCGCTGG	388
651	AGGGGAGACTGACGCGCAAGGAGGGCCGAAAGTTCTTCGCGAGATTTGTGTTCGCGCTGG	710
389	ACTTCTGCCACAGCCACTCCATATGCGCACAGGGATCTGAAAACCTGAAAAACCTCTCTCTGG	448
711	ACTTCTGCCACAGCTACTCTCATCTGCGCACAGAGACTTAAAGCCCGAGAACCTGCTTTTGG	770
449	ACGAGAAGAACAACTCCGCATCGCAGACTTTGGCATGGCGTCCCTGCAAGTGTGGCGACA	508
771	ATGAGAAAAAACATCCGCAATTGCAGACTTCGGCATGGCGTCCCTGTCAGGTGGGGGACA	830
509	GCCTGTTGAGAGACGAGTGTGGTTCGCCCACTACGCCTGCCCGGAGAGGTGATCCGGGGG	568
831	GCCTCTCTGAGAGCAGCTCGGGTCCCGCCCATTTATGCGTGTCCAGAGGTGATTTAAGGGGG	890
569	AGAAGTATACAGCGCGGAAGGGGAGCTGTGAGACTGCGCGCTCATCTTGTTCGCTTGC	628
891	AAAAATATGATGCGCCCGGGGAGACATGTGAGCTGTGAGTCACTCTTTCGCGCTGC	950
629	TGTTGGGGGCTCTGCCCTTTCGAGATGACAACTTGCGACAGCTGTCTGGAGAAAGGTGAAGC	688
951	TCGTGGGGGCTCTGCCCTTTGATGACGACAACTCCGCGAGCTGCTGGAGAGGTGAAC	1010
689	GGGGCGTTTCCATGTCGCGACTTTATCCGCGCCACTGCGCAGAGTCTGCTAGGGGGCA	748
1011	GGGGCGCTTTCCACATGCCCCACTTCAITTCCTCAGATTGCCAGAGCCCTCTCAGAGGGAA	1070
749	TGATCAGAGTGGAGCGCGGACCGGCTCAGCTAGAGCAGCATTTCAGAAACACATATGTT	808
1071	TGATCAGAGTGGAGCGCGGAAAGAGCTGATCTGGAGCAATTCAGAAACATCTCTTGT	1130
809	ATATAGGGGGCAAGATGAGCCCGAAACC-----AGAGCAGCCCATTTCTTCGCAAGTGC	862
1131	ACCTAGCGGGAAACACGAGCCAGACCGGTGCTGGAGCCAGCCCTGGCGCGCGGTAG	1190
863	AGATCGCTCGTGCACAGCCTTGAGAGACATCGACCCCGAGCTGTCTGGACAGGATGCAC	922
1191	CCATCGGGAGCCTGCCATCCAAACGAGAGCTGACACCCGACGCTCTAGAGAGCATGGCAT	1250
923	CATCTGGGCTCTTCCGAGACCGCAACAGCTGTGAGGACCTGCTCTCGAGAGGAGAGA	982
1251	CATCTGGGCTCTTTCAGGACCGGAGAGGCTGATCTCGGAGCTGCGCAGTGTGAGAGGAGA	1310
983	ACCAAGAGAAAGATGATTTACTTCTCTCTGACCGGAAAGAAAGTATCCCGAGCCAGG	1042
1311	ACCAAGAAAGATGATATATATCTGCTTTTGATCGGAGAGAGCGGTATCCCGAGCTGT	1370
1043	AGGATAGGACCTTCCCGCCCGGAAACGAGATAGACCTTCCCGGGAACGGTGTGTGACTCCC	1102
1371	AGGACCAAGGACCTGCCTCCCGGAATGATGTTGACCCCCCGGGAACGGTGTGTGATCTC	1430
1103	CGATGCTGAACCGGCACGGCAAGGGCGGCAGAACGCAATCCATGGAGGTGCTCAGCG	1162
1431	CCATGCTGAGCCGTCTACGGGAAGCGGCAACGAGCGGAAGTCCATGGAATCTCTGAGCA	1490

```
QY 1163 TGAC-----GGACGGCGGCTCCCGGTGCTTCCCGCGGGGCCATTGAGATGGCCC 1213
Db 1491 TCACCGATGCCGGGGTGGTCCCTCTGATCCACCCGACGGGCTTGAGATGGCCC 1550
QY 1214 AGCAGGCGCAGAGTCTCGGTCCATCAGCGGTGCTCTCAGGGCTTTTCACACAGCCAC 1273
Db 1551 AGCAGGCGCAGAGTCCCGGTAGCGTCACTGAGCGCTCCACGGGTCTGTCTCCAGCCCTC 1610
QY 1274 TCAGCAGCCCGCGGT-----ACCCCAAGGGGACACCTCTCCACACGCCAA 1357
Db 1611 TAGCAGCCCAAGAGTCCGGTCTTTTCTTTTCCACCGGACCGGGGCTGGAGATGAGG 1670
QY 1290 -----GACCCCTACCCCTCACCAGGGGAGTCCCTCCCTCC----- 1326
Db 1731 GTGGGGCGCGGGAGCAGCCCGCCCGCCCGCTCAGTCCCGCTCCACACCCCTGCGCGCC 1790
QY 1327 -----ACCCCAAGGGGACACCTCTCCACACGCCAA 1357
Db 1791 CCCAGGCTCCCGGCTCTCTGGCGGACCCCTTGCATCTCGCTCTGCACACGCCCC 1850
QY 1358 AGGAGAGCCCGGCTGGCACGGCCCAACCCACGCCCCCGCTCCAGCCCC-----AGCGTCG 1411
Db 1851 GGGCCAGTCCACCGGGACCCCGGGGACAAACACACCCCGCCCGGGCGGTGGGTGCG 1910
QY 1412 GAGGGTCCCTGGAGGGCGGCTCACTCCATCAGAACAGCTTTCTGGGCTACGCC 1471
Db 1911 GGGAGGCGCTGGAGGAGTCTCTCACTCCATCCGCAACAGCTTCTGGGCTCCCTC 1970
QY 1472 GCTTCCACCGCGGAAACTGCAAGTTCCGACGCGCGGAGGATGTCCAACTGACACAG 1531
Db 1971 GCTTTCACGGCGAGATGAGGTCTTACCGCTGAGGAGATGTCCAGCTTGAGCGCAG 2030
QY 1532 AGTCGTCCCGAGAGTGGCGAAGAGTCTCGTTTGGGAACTTCATCAGCCTGGAGAAGG 1591
Db 2031 AGTCCTCCCGAGTGGCAAAAAGCTCTCTGTTTGGGAACTTCATCTCTTGGACAAAG 2090
QY 1592 AGGAGCAGATCTTCTGTGTCATCAAGACAAACCTCTGAGCTCCATCAGCTGACATCG 1651
Db 2091 AAGAACAAATATCTCTGTGTCATCAAGGACAAACCTCTCAGCAGCATCAAGACACATCG 2150
QY 1652 TGCAGCCTCTCTGTGATTCCTAGTCTCAGCTCAGCAGCTCATCTCCCAACAGCTTCC 1711
Db 2151 TCCATGCTTTCTGTGATTCCTAGCTCAGCTCAGCTCAGCTGCTGTCAAGACACAGCTTCA 2210
QY 1712 GGGCCGAGTACAAGGCCACGGGGGGCGGCGGCTGTTCAGAACCGGCTCAAGTTCCAGG 1771
Db 2211 GGGCCGAGTACAAGGCCAGTGGCGGCGCTCCGCTCTTCCAAAGCCGCTCCGCTTCCAGG 2270
QY 1772 TTGATATCACTACACGAGGCT-----GGGAGGCGCGAGAGGAGA 1813
Db 2271 TGGACATCAGTCTCTGAGGCTCAGAGCCCTCCCGCGACGGGACGGCAGCGAGGTG 2330
QY 1814 ACGGCATCTACTCGTCACTTCACTCTGCTCAGGCGCCCGGCGCTTCAAGAGG 1873
Db 2331 GTGGCATCTACTCGTCACTTCACTCTGCTCAGGCTCCAGCGCTTCAAGCGAG 2390
QY 1874 TGGTGGAGACCATCCAGGCCAGTGTGTGAGCACACACGACCCGCTCGCGCCCGAGCACT 1933
Db 2391 TGGTGGAGACCATCCAGGCCAGTGTGTGAGCACACGACCCGCTCGCGCCCGAGCACT 2450
QY 1934 TGTGAGACACACTTAATG 1952
Db 2451 TGGCAGACGAGAAGAACGG 2469
```

RESULT 11
AX166526
LOCUS

AX166526 2395 bp DNA linear PAT 22-JUN-2001

```
DEFINITION Sequence 17 from Patent WO0138503.
ACCESSION AX166526
VERSION AX166526.1 GI:14546871
KEYWORDS Homo sapiens (human)
SOURCE Homo sapiens
ORGANISM Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1
AUTHORS Plowman,G.D., Whyte,D., Manning,G.S., Sudarsanam,S.S., Martinez,R.,
Flanagan,P. and Clary,D.S.
TITLE Novel human protein kinases and protein kinase-like enzymes
JOURNAL Patent: WO 0138503-A 17 31-MAY-2001;
Sugen, Inc. (US)
FEATURES Location/Qualifiers
source 1..2385
/mol_type="genomic DNA"
/db_xref="taxon:9606"
BASE COUNT 494 a 768 c 731 g 392 t
ORIGIN
Query Match 48.1%; Score 964.4; DB 6; Length 2385;
Best Local Similarity 69.8%; Pred. No. 2.9e-162;
Matches 1490; Conservative 0; Mismatches 451; Indels 195; Gaps 6;
QY 12 GGGGAAGGACGGCGCGCAGCACGCGAGTATGTTGGCCCTACCGGCTGGAGAAGAC 71
Db 105 GGAGGCGAAGAGAGAGGGCGCCACCCCATATGTTGGCCCTATCGGCTGGAGAAGAC 164
QY 72 GCTGGGCAAGGCGCAGACAGCTGCTGGTGAAGCTGGGGGTTCACTGGCTCACTGCCAGAA 131
Db 165 GCTGGGCAAGGCGCAGACAGCTGCTGGTGAAGCTGGGGTCACTGGCTCACTGCCAGAA 224
QY 132 GGTGGCCATCAAGATCGTCAACCGTGAGAGCTACGCGAGTCTGCTGATGAGGTGA 191
Db 225 GGTGGCCATCAAGATCGTCAACCGTGAGAGCTACGCGAGTCTGCTGATGAGGTGA 284
QY 192 GCGGGAGATCGGCATCTGAAGCTCAATGAGACACCCAGTCTTAAAGCTGACACGCT 251
Db 285 GCGGGAGATCGGCATCTGAAGCTCAATGAGACACCCAGTCTTAAAGCTGACACGCT 344
QY 252 TTATGAAAACAAAATATTTGTACCTGTGTGTAGAACACAGTGTCAAGTGGTGAAGCTCTT 311
Db 345 CTACGAGAACCAAGATATTTGTACCTGTGTGTAGAACACAGTGTCAAGTGGTGAAGCTCTT 404
QY 312 CGATACCTGTGTGAAGAGGGAGGCTGACCCCTAAGAGGCTCGGAAGTCTTCCGGCA 371
Db 405 CGATACCTGTGTGAAGAGGGAGGCTGACCCCTAAGAGGCTCGGAAGTCTTCCGGCA 464
QY 372 GATCATCTCTGGCTGAGCTTTGCGCACAGCCACTCCATATGCCAGGGATCTGAAACC 431
Db 465 GATGTGTCTGGCTGAGCTTTGCGCACAGCTTCTGCGCACAGCTTCTGCGCACAGGACCTTAAGCC 524
QY 432 TGAAAACCTCTCTGTGGAGAGAGAACCAACATCCGCATCGCAGACTTTGGGATGGCGTC 491
Db 525 CGAAGACCTCTCTGTGGAGAGAGAACCAACATCCGCATCGCAGACTTTGGGATGGCGTC 584
QY 492 CTTGAGGTTGGCGACAGCTTTGAGAGACAGCTGTGGGTCCCCCACTACCGCTTCCGCC 551
Db 585 CTTGAGGTTGGCGACAGCTTTGAGAGACAGCTGTGGGTCCCCCACTACCGCTTCCGCC 644
QY 552 CGAGGTGATCCGGGGGAGAGATATGACCGCGGAGAGGCGGCTGTGGAGCTGGCGGT 611
Db 645 AGAGGTGATTAAGGGGAGAGATATATATGTCGCCCGCGGAGACATGTGGAGCTGTGGAGT 704
QY 612 CATCTCTTTCGCTTGTGTGGGGCTCTGCCCTTTCGACGATGACAACTTGGACAGCT 671
Db 705 CATCTCTTTCGCTTGTGTGGGGCTCTGCCCTTTCGACGATGACAACTTGGACAGCT 764
QY 672 GCTGGAGAGGTGAAGGGGGCGTGTTCACATGCGGCACTTTATCCCGCCGACGTGCCA 731
Db 765 GCTGGAGAGGTGAAGGGGGCGTGTTCACATGCGGCACTTTATCCCGCCGACGTGCCA 824
```

```
QY 732 GAGTCTCTAGCGGGCATGATCGAGGTGGAGCGCGACCGCCTCACGCTAGAGCACAT 791
Db 825 GAGCTCTCTGAGGGGATGATCGAGTGGAGCGCCGAAAAAGAGCTAGTCTGAGGCAAT 884
QY 792 TCAGAAACACATATGGTATATAGGGGCAAGATGAGCCCGAAACC-----AGAGCAGCC 845
Db 885 TCAGAAACATCTCTGGTACCTTAGCGGGGAAACAGAGCCAGACCCGTCGCTGGAGCAGC 944
QY 846 CATTCCTCGAAGGTGAGATCGCTCGCTGCCAGCGCTCGAGGACATCGACCCCGACGT 905
Db 945 CCCTGGCGCGGGTAGCCATGCGGAGCGCTGCCATCAACGAGAGAGTGGACCCCGACGT 1004
QY 906 GCTGGAACAGCATGCACTCACTGGGCTCTTCCGAGACCGCAACAAAGCTGTGCGAGACCT 965
Db 1005 CTTAGAGAGCATGCACTCACTGGGCTCTTCCGAGACCGCGAGAGGCTGCTCGGAGCT 1064
QY 966 GCTGTCGAGAGAGAGAACACAGGAGAGATGATTTACTTCTCTCTGAGACCGAAAGA 1025
Db 1065 GCGCAGTGGAGAGAGAGAACCAAGAAAGATGATATATTATCTGCTTTTGGATCGGAAGGA 1124
QY 1026 AAGGTACCCGAGCCAGGAGATGAGGACCTGCCCGCCCGGAAAGAGATAGACCTCCCGG 1085
Db 1125 GCGGTATCCCACTGTGAGGACAGGACCTGCCCTCCCGGAATGATTTGACCCCGCCCG 1184
QY 1086 GAAGCGTGTGAGTCTCCCGATGCTGAACCGGCAAGCGGCGCGAGAAACGCAATC 1145
Db 1185 GAAGCGTGTGATTTCTCCATGCTGAGCGGTCAAGGAGCGGACCGAGCGGAAGTC 1244
QY 1146 CATGAGAGTGTCTAGCGTGAC-----GGAGCGGGCTCCCGGTGCTGGCGGGG 1196
Db 1245 CATGGAAGTCTGAGCATCACCGATGCGGGGGTGGTGGCTTCCCTGTACCCACCGGACG 1304
QY 1197 GGCAATTGAGATGCCAGACGCGCAGAGGTCTCGGTTCATCAGCGGTGCTCTCTCAGG 1256
Db 1305 GGCCTTGAGATGGCCACAGCACAGCAGAGATCCGTAGCGTCTAGTGAGGCTTCCAGGG 1364
QY 1257 CCTTTCACACGCGCTCTAGCAGCGCCCGGGT----- 1289
Db 1365 TCTGTCTCCAGCGCTCTAAGCAGCCCAAGAGTCCGGTCTTTCTCTTTTTCACGGAGCC 1424
QY 1290 ----- 1289
Db 1425 GGGGGTGGAGATGAGGCTCGAGGGGGGGCTCCCGACTTCCAAACGAGAGCGTGGCC 1484
QY 1290 -----GACCCCTCACCCCTCACCAAGGGG 1313
Db 1485 TTCTCGGGGCCCCAGGGTGGGGGGCGGGGAGCAGCCCCCGGCCAGTGGCGCTC 1544
QY 1314 CAGTCCCTCTCCC-----ACCCCAAGGGGAC 1340
Db 1545 CACACCCCTGCGCGCGCCCCCAGGCTCCCTGCGGGGAGCCCGCTTGCACCTC 1604
QY 1341 ACCTGTCCACAGCCAAAGGAGAGCGCGTGGCAGCGCCCAACCCAGCGCCCGTCCAG 1400
Db 1605 GCCTTGGACACGCGCCCGGCGCAGTCCACCGGACCCCGGGGACAAACACACCCCGCAG 1664
QY 1401 CCCC-----AGCCTCGAGGGGTGCTCTGAGGGGCGGGTCAACTCTCCATCAAGAACAG 1454
Db 1665 CCCCAGGGTGGCTCGGGGGAGCCGCTTGGAGAGTCTCAACTCTCATCCGCAACAG 1724
QY 1455 CTTTCTGGGTCTCA CCGCTTTCACCGCGGAAACTGCAAGTTCCGACGCGCGAGGAGAT 1514
Db 1725 CTTCTGGGTCTCCCTCGCTTTTCAACCGCGCAAGATGACGAGTCTTCCGCTGAGGAGAT 1784
QY 1515 GTCCAACTGACACAGAGTCTCCAGAGCTCGCGAGAGAGTCTCTGTTTGGGAATCT 1574
Db 1785 GTCCAGCTTGAAGCGCAGAGTCTCTCCCGGAGCTGGCAAAACGCTCTGGTTCGGGAATCT 1844
QY 1575 CATCAGCTCGAGAGGAGGAGAGATCTCTCGTGGTCTATCAAGACAAACCTCTGAGCTC 1634
Db 1845 CATCTCTTGGACAAAGAGAACAAATATTCTCTGTGTAAAGGACAAACCTCTCAGCAG 1904
```

```
QY 1635 CATCAAGGCTGACATCGTCGACGCGCTTCTCTGATTCCTCCAGTCTTCAGCCACAGCGTAT 1694
Db 1905 CATCAAGACAGACATCGTCCATGCTTCTGTCGATCCCGACCTGAGTCAAGTGTGCT 1964
QY 1695 CTTCCAAACGAGCTTCCGGGCGGAGTACAGCCACGGGGGGGCGAGCGTGTTCAGAA 1754
Db 1965 GTCAACAGACCAAGCTTCAAGGGCGGAGTACAAGGCCAGTGGCGGCGCTCTCGTCTTCCAAA 2024
QY 1755 GCGGTCTCAAGTCTCCAGGTTGATATCACTACACGAGGCT-----GG 1796
Db 2025 GCGCTCGCTTCCAGTGGGACATCAGCTCTCTGAGGGTCCAGAGCCCTCCCGCGAGC 2084
QY 1797 GGAGCGCAGAAAGAGAACGGCATCTACTTCGTCACCTTCACTCTCAGGCGCCAG 1856
Db 2085 GGACGCGCAGGAGGTGGTGGCATCTACTTCGTCACCTTCACTCTCATCTCGGGTCCAG 2144
QY 1857 CGCTCGCTCAAGAGGGTGGTGGAGACCATCAGCGCCAGCTGCTGAGCACACGACCC 1916
Db 2145 CGCTCGGTTCAGCGAGTGGTGGAGACCATCAGGACAGCTCTCTGAGCACTCATGACCA 2204
QY 1917 GCGTGGCGCCAGCAGCTTGTTCAGACACCACTAACTG 1952
Db 2205 GCGCTCGGTGAGGCGCTGGCAGACGAGAGACGG 2240

RESULT 12
AX642966
LOCUS 2897 bp DNA linear PAT 24-FEB-2003
DEFINITION Sequence 43 from Patent WO01096547.
ACCESSION AX642966
VERSION AX642966.1 GI:29550113
KEYWORDS Homo sapiens (human)
SOURCE Homo sapiens
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1
AUTHORS Yue,H., Lal,P., Bandman,O., Borowsky,M., Au-Young,J., Lu,Y.,
Gandhi,A.R., Tribouley,C.M., Walla,N., Yao,M.G., Lu,D.A.,
Greenwald,S.B., Ramkumar,J., Griffin,J.A., Kearney,L., Burford,N.,
Nguyen,D.B., Tang,Y.T., Baughn,M.R., He,A., Thornton,M.,
Hafalia,A., Patterson,C., Gururajan,R., Lo,T.P., Khan,F.,
Recipon,S.A., Azimzal,Y., Pollick,J.L., Ding,L., Grether,M.,
Elliot,V.S., Thangavelu,K., Batra,S. and Ison,C.H.
HUMAN KINASES
TITLE Incyte Genomics, Inc. (US)
JOURNAL Patent: WO 01096547-A 43 20-DEC-2001;
Incye Genomics, Inc. (US)
FEATURES
Source 1..2897
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/notes="Incyte ID No: 7477486CB1"
BASE COUNT 599 a 921 c 877 g 500 t
ORIGIN
Query Match 48.1%; Score 964.4; DB 6; Length 2897;
Best Local Similarity 69.8%; Pred. No. 2.8e-162;
Matches 1490; Conservative 0; Mismatches 451; Indels 195; Gaps 6;
QY 12 GGGGAGGAGCGGGCGCGCAGCAGCGCGCTATGTTGGCGCTTACCGGCTGGAGAGAC 71
Db 105 GGAGCGCGAAGAGAGAGAGGCGCGCCCAATATGTGGCGCTTATCGGCTGGAGAGAC 164
QY 72 GCTGGGCAAGGGGCGACAGAGGTCTGGTGAAGCTGGGGGTTCACTGCTCACCTGCCAGAA 131
Db 165 GCTGGGCAAGAGACAGAGAGGCTGGTTAACTCGGGGTCCTGATCATCGGCTCAGAA 224
QY 132 GTTGGCCATCAAGATCGTCAACCGTGAGAGCTCAGCGAGTGGTCTGATGAAGTGA 191
Db 225 GGTGCGCATCAAGATCGTGAACCGGAGAGCTGTGAGAGTGGTCTGATGAAGTGA 284
QY 192 GCGGAGATCGCGATCTGAAGCTCATTTAGGACACCCCGCTTAAAGCTGCACGACGT 251
```


TITLE Direct Submission
Submitted (02-FEB-2002) School of Life Sciences, Laboratory of
Human Genes Research, Institute of Genetics, Fudan University, 220
Handan Road, Shanghai 200433, P. R. China

FEATURES

source Location/Qualifiers
1.3007
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/chromosome="19"
/map="19q13"
/tissue_type="brain"
128..2512
CDS
/codon_start=1
/product="putative serine/threonine protein kinase"
/protein_id="AAL87697.1"
/db_xref="GI:19401871"
/translation="MWAGLILGKPSPDGVSVPKRDVAGGGGEEAEERGRHA
QYVGYRLEKLTGKGTGLVKGVCITGQKVAIKVNRKLSVLMKVEREIALIK
LI EHPVHLKLDVYENKYLVLVLEHVSQGLFDYLVKGRITPKEARFFQIVASAL
DFCHSYISICHRDLKPNLLDEKNIRIADFGWASLQVDSLLETSCGSPHYACPEVI
KGRKYDGRADMSQVILLFALLVGLPDDNNLRLLLEKVRGVFHPHFIPDPCV
LLRGMIEVEREKLSLEIQIKHPWYLGKHEPDPCLPAPARRVNRSLPSNGELDPD
VLESMSLGCFRDRRLRRLRERKESEENQEMIIYLLDRKERYPSCEDQDLPPRNDV
PPRRYDPSMLSRHGRRLRERKSMELVLSITDAGGSGSPVPTRALEMAHQSRSDVS
GASTGLSSPLSPSPVFSPEPGADGDEARGSGPTKTQTLPRSPGPGGAGAPR
PPSARSPTLPSPGPPSPRPLHPLHTPRASPTGPTTLPSPGPGVGAAR
SRLSNRNSPLGSPRPHRRKQVPTAEEMSLTPRESSPELAKSFGNFIISLDRQI
FLVLDKPLSSIRADIVHAFLSIPSLSHSVLSOTSRAEYKASGSPSVQKPRFQVD
ISSEGEPSPRNDGGGSIYSITFLISGPRFRFRVETIQALLSLTHDQPSVQA
LADEKGAQTRPAGAPRPSLPQPPRPDPELSSRRPGRPKKLLATNGTLP"

BASE COUNT
ORIGIN

636 a 940 c 906 g 525 t

Query Match 48.1%; Score 964.4; DB 9; Length 3007;

Best Local Similarity 65.8%; Pred. No. 2.8e-162;
Matches 1490; Conservative 0; Mismatches 451; Indels 195; Gaps 6;

Qy 12 GGGGAAGAGCGCGCGCGAGCAGCGCGCAGTATGTTGGGCCCTTACCGCTGGAGAAGAC 71
Db 232 GGAGCGCGAAGAGAGAGGGCGCCACGCCAATATGTGGGCCCTATCGGCTGGAGAAGAC 291
Qy 72 GCTGGCAAGGGGAGAGAGCTCTGTTGAAGCTGGGGGTTCACTGCGTCACTTCCAGAA 131
Db 292 GCTGGCAAGAGAGAGAGCGGCTGGTTAACTCGGGGTCCACTGCATCACGGGTCAAA 351
Qy 132 GGTGGCCATCAAGATCTCAACCTGAGAGAGCTCAGCGAGTCGCTGATGAGGTGA 191
Db 352 GGTGGCCATCAAGATCTGTAACCGGGAGAGCTGTCCGAGTCGCTGATGAAAGGTGA 411
Qy 192 GCGGAGATCGGATCTCTGAAGCTCATTTGAGCAGCCCGCCAGCTCTCAAGCTGCACGACGT 251
Db 412 GCGGAGATCGGATCTCTGAAGCTCATCGACAGCCAGCTCTCAAGCTGCACGACGT 471
Qy 252 TTATGAAGCAAGAAATATTTGTACTGTGCTGCTAGAAACAGTGTGAGTGTGAGTCTTT 311
Db 472 CTAGCAGAAACAAGAAATATTTGTACTGTGCTGAGCAGCTCTCGGGGGGTGAGCTATT 531
Qy 312 CGACTACTGTGTAAGAGGGAGGCTGACCGCTTAAGAGGAGCTCGGAAGTCTTCCCGCA 371
Db 532 CGACTACTGTGTAAGAGGGAGAGCTGACCGCCCAAGAGGGAGGCTGCTTCCCGCA 591
Qy 372 GATCATCTCTGCGTGGATCTCTGCCACAGCACTTCCATATGCCACAGGATCTGAAACC 431
Db 592 GATTGTGCTGCGCTGGACTTCTGCCACAGCTACTCCATCTGCCACAGAGACCTAAAGCC 651
Qy 432 TGAAGAACTCTGCTGGAGAGAGAAACAATCCGCAATCCGAGACTTTGGCATGGCGTC 491
Db 652 CGAGAACCTGCTTTTGGATGAGAAAAACAACATCCGCAATTCAGACTTTCCGATGGCGTC 711
Qy 492 CCTGAGGTTCGCAAGCTGTTGGAGACAGCTGTGGGTCCCCCACTACGCTGCC 551
Db 712 CTTGCAAGTGGGGAGAGCTCTCTGGAGACAGCTGGCGGTCCCCCACTATGCTGTCC 771

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 27, 2003, 09:04:30 ; Search time 1165.25 Seconds
(without alignments)
5674.682 Million cell updates/sec

Title: US-10-054-579-1

Perfect score: 2007

Sequence: 1 atgacatcgacgggaagga.....gaattatccgaaagttaa 2007

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 2190069 seqs, 1647345023 residues

Total number of hits satisfying chosen parameters: 4380138

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications NA:*

- 1: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq:*
- 2: /cgn2_6/ptodata/2/pubpna/PCT_NEW_PUB.seq:*
- 3: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq:*
- 4: /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq:*
- 5: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq:*
- 6: /cgn2_6/ptodata/2/pubpna/PCTUS_PUBCOMB.seq:*
- 7: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq:*
- 8: /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq:*
- 9: /cgn2_6/ptodata/2/pubpna/US09A_PUBCOMB.seq:*
- 10: /cgn2_6/ptodata/2/pubpna/US09B_PUBCOMB.seq:*
- 11: /cgn2_6/ptodata/2/pubpna/US09C_PUBCOMB.seq:*
- 12: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq:*
- 13: /cgn2_6/ptodata/2/pubpna/US10A_PUBCOMB.seq:*
- 14: /cgn2_6/ptodata/2/pubpna/US10B_PUBCOMB.seq:*
- 15: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq:*
- 16: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq:*
- 17: /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2007	100.0	2007	13	US-10-054-579-1
2	2005.4	99.9	2908	14	US-10-195-072-1
3	2005.4	99.9	2908	14	US-10-195-071-1
4	1993	99.3	3364	14	US-10-195-072-3
5	1993	99.3	3364	14	US-10-195-071-3
6	1966	98.0	2647	12	US-10-288-798-42
7	1941.2	96.7	2025	14	US-10-283-247-1
8	1939.6	96.6	2025	10	US-09-842-582-3
9	1939.6	96.6	2219	10	US-09-842-582-1
10	1827	91.0	1827	13	US-10-054-579-3
11	1685.2	84.0	1911	14	US-10-283-247-4
12	965	48.1	2337	12	US-10-116-326-1
13	965	48.1	2980	12	US-10-354-358-91
14	851.4	42.4	2112	12	US-10-116-326-5
15	777.2	38.7	2289	12	US-10-116-326-3
16	594	29.6	2173	12	US-10-102-548-1

17	594	29.6	2173	14	US-10-102-558-1	Sequence 1, Appli
18	291.6	14.5	512	10	US-09-960-253-87	Sequence 87, Appl
19	265.8	13.2	3609	12	US-10-161-565-19	Sequence 19, Appl
20	265	13.2	1594	14	US-10-284-060-1	Sequence 1, Appli
21	265	13.2	1594	14	US-10-284-060-3	Sequence 3, Appli
22	264.2	13.2	2085	12	US-10-161-565-20	Sequence 20, Appl
23	264.2	13.2	2278	12	US-10-161-565-21	Sequence 21, Appl
24	264.2	13.2	3226	12	US-10-161-565-23	Sequence 23, Appl
25	264.2	13.2	4917	12	US-10-161-565-22	Sequence 22, Appl
26	262.6	13.1	3270	10	US-09-835-081-1	Sequence 1, Appli
27	252.4	12.6	2352	9	US-09-815-915-3	Sequence 3, Appli
28	252.4	12.6	2352	12	US-10-393-316-3	Sequence 3, Appli
29	252.4	12.6	2968	9	US-09-815-915-1	Sequence 1, Appli
30	252.4	12.6	2968	12	US-10-393-316-1	Sequence 1, Appli
31	229.4	11.4	3154	10	US-09-919-585-19	Sequence 19, Appl
32	223.4	11.1	2349	11	US-09-823-187-25	Sequence 25, Appl
33	214.6	10.7	1539	10	US-09-938-842A-2657	Sequence 2657, Ap
34	214.6	10.7	3609	9	US-09-799-875-6	Sequence 6, Appli
35	214.6	10.7	4381	12	US-10-288-798-31	Sequence 31, Appl
36	214.6	10.7	5983	9	US-09-799-875-4	Sequence 4, Appli
37	214.2	10.7	2086	12	US-10-161-565-4	Sequence 4, Appli
38	214.2	10.7	2112	10	US-09-919-585-4	Sequence 4, Appli
39	214.2	10.7	2112	12	US-10-161-565-7	Sequence 7, Appli
40	214.2	10.7	2248	10	US-09-919-585-7	Sequence 7, Appli
41	214.2	10.7	2248	12	US-10-161-565-5	Sequence 5, Appli
42	214.2	10.7	2701	12	US-10-161-565-6	Sequence 6, Appli
43	214.2	10.7	2784	12	US-10-161-565-2	Sequence 2, Appli
44	214.2	10.7	2946	12	US-10-161-565-1	Sequence 1, Appli
45	214.2	10.7	3103	12	US-10-161-565-3	Sequence 3, Appli

ALIGNMENTS

RESULT 1

US-10-054-579-1

; Sequence 1, Application US/10054579

; Publication No. US20020137913A1

; GENERAL INFORMATION:

; APPLICANT: Turner, C. Alexander Jr.

; APPLICANT: Mathur, Brian

; TITLE OF INVENTION: No. US20020137913A1 Human Kinases and Polynucleotides Encoding t

; FILE REFERENCE: LEX-0300-USA

; CURRENT APPLICATION NUMBER: US/10/054,579

; CURRENT FILING DATE: 2002-01-22

; PRIOR APPLICATION NUMBER: US 60/263,378

; PRIOR FILING DATE: 2001-01-23

; NUMBER OF SEQ ID NOS: 4

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 1

; LENGTH: 2007

; TYPE: DNA

; ORGANISM: homo sapiens

; US-10-054-579-1

Query Match 100.0%; Score 2007; DB 13; Length 2007;

Best Local Similarity 100.0%; Pred No. 0;

Matches 2007; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy	1	ATGACATCGACGGGGAAGGACGCGCGCGCAGCAGCAGCGGCGCTATGTTGGCCCTACCGG	60
Db	1	ATGACATCGACGGGGAAGGACGCGCGCGCAGCAGCAGCGGCGCTATGTTGGCCCTACCGG	60
Qy	61	CTGGAGAAGCGCTGGGCAAGGGGAGGAGCGTCTGTTGAGCTGGGGGTTCTACTGGCTC	120
Db	61	CTGGAGAAGCGCTGGGCAAGGGGAGGAGCGTCTGTTGAGCTGGGGGTTCTACTGGCTC	120
Qy	121	ACCTGCCAGAGGTGGCCATCAAGATCGTCAACCGTGAGAGCTCAGCGAGTCGGTCTG	180
Db	121	ACCTGCCAGAGGTGGCCATCAAGATCGTCAACCGTGAGAGCTCAGCGAGTCGGTCTG	180
Qy	181	ATGAAGTGGAGCGGGAGATCGCGATCTCGAAGCTCATTTAGCACCACCCAGCTCTAAAG	240
Db	181	ATGAAGTGGAGCGGGAGATCGCGATCTCGAAGCTCATTTAGCACCACCCAGCTCTAAAG	240

Query Match	99.9%;	Score	2005.4;	DB	14;	Length	2908;	
Best Local Similarity	100.0%;	Pred. No.	0;					
Matches 2006;	Conservative	0;	Mismatches	1;	Indels	0;	Gaps	0;
QY	1	ATGACATCAGCGGGGAAGGACGGCGCGCGAGCAGCGAGTATGTTGGGCCCTACCGG	60					
DB	106	ATGACATCAGCGGGGAGAGACGGCGCGCGAGCAGCGAGTATGTTGGGCCCTACCGG	165					
QY	61	CTGGAGAAGACGCTGGGCAAGGGGAGAGACAGTCTGTGTAAGCTGGGGGTTCACTGCGTC	120					
DB	166	CTGGAGAAGACGCTGGGCAAGGGGAGAGACAGTCTGTGTGAGCTGGGGGTTCACTGCGTC	225					
QY	121	ACCTGCCAAGAGTGGCCATCAAGATCGTCAACCGTGAAGAGCTCAGGAGTGGGTGCTG	180					
DB	226	ACCTGCCAAGAGTGGCCATCAAGATCGTCAACCGTGAAGAGCTCAGGAGTGGGTGCTG	285					
QY	181	ATGAAGGTGGAGCGGAGATCCGATCTCTGAAGCTCATTTGAGCACCCGCCAGCTCTAAAG	240					
DB	286	ATGAAGGTGGAGCGGAGATCCGATCTCTGAAGCTCATTTGAGCACCCGCCAGCTCTAAAG	345					
QY	241	CTGCACAGCTTTATGAAGACAAAATAATTTGTACTCTGTCTAGAACACGTCTCAGGT	300					
DB	346	CTGCACAGCTTTATGAAGACAAAATAATTTGTACTCTGTCTAGAACACGTCTCAGGT	405					
QY	301	GGTGAGCTCTTGAATCTGCTGTGAAGAGGGAGGCTGACGCCCTAAGGAGGCTCGGAAG	360					
DB	406	GGTGAGCTCTTGAATCTGCTGTGAAGAGGGAGGCTGACGCCCTAAGGAGGCTCGGAAG	465					
QY	361	TTCTTCCGCGAGATCATCTCGGCTGGACTTCTGCGACAGGCACTCCATATGCCACAGG	420					
DB	466	TTCTTCCGCGAGATCATCTCGGCTGGACTTCTGCGACAGGCACTCCATATGCCACAGG	525					
QY	421	GATCTGAACCTGAAACCTCTGCTGGACGAGAAACAAATCCGATCCGATCGCAGACTTT	480					
DB	526	GATCTGAACCTGAAACCTCTGCTGGACGAGAAACAAATCCGATCCGATCGCAGACTTT	585					
QY	481	GGCATGGCTCCCTGAGGTTGGGACAGCTCTGTTGGAGACAGCTGTGGGTCCCCCAC	540					
DB	586	GGCATGGCTCCCTGAGGTTGGGACAGCTCTGTTGGAGACAGCTGTGGGTCCCCCAC	645					
QY	541	TAGCGCTCCCGAGGTGATCCGGGGGAGAGTATGACGGCCGAGGCGGAGCTGTGG	600					
DB	646	TAGCGCTCCCGAGGTGATCCGGGGGAGAGTATGACGGCCGAGGCGGAGCTGTGG	705					
QY	601	AGCTGCGGGCTCATCTGCTTGCCTTGTGTGGGGGCTCTGCCCTTCGACGATGACAAAC	660					
DB	706	AGCTGCGGGCTCATCTGCTTGCCTTGTGTGGGGGCTCTGCCCTTCGACGATGACAAAC	765					
QY	661	TTGCGACAGCTGCTGGAGAAGGTGAAGCGGGCGTGTTCCACATGCCGCACTTTATCCCG	720					
DB	766	TTGCGACAGCTGCTGGAGAAGGTGAAGCGGGCGTGTTCCACATGCCGCACTTTATCCCG	825					
QY	721	CCGCACTGCCAGTCTGCTACGGGGCATGATCGAGGTGGACCGCGCAGCGGCTCACG	780					
DB	826	CCGCACTGCCAGTCTGCTACGGGGCATGATCGAGGTGGACCGCGCAGCGGCTCACG	885					
QY	781	CTAGACACATTCAGAAAAACATATGGTATATAGGGGCAAGATGAGCCGAAACAGAG	840					
DB	886	CTAGACACATTCAGAAAAACATATGGTATATAGGGGCAAGATGAGCCGAAACAGAG	945					
QY	841	CAGCCCATTCCTCGCAAGGTGCAGATCCGCTCGCTGCCAGCTTGGAGGACATCGACCCC	900					
DB	946	CAGCCCATTCCTCGCAAGGTGCAGATCCGCTCGCTGCCAGCTTGGAGGACATCGACCCC	1005					
QY	901	GAGTGTGAGACAGCATGCACTCATGTGGGCTGTTCCGAGACCGCAACAGCTGTGAG	960					
DB	1006	GAGTGTGAGACAGCATGCACTCATGTGGGCTGTTCCGAGACCGCAACAGCTGTGAG	1065					
QY	961	GACCTGCTGCCAGGAGGAGACAGGAGAGATGATTACTTCTCTCTCTCTGACCGG	1020					
DB	1066	GACCTGCTGCCAGGAGGAGACAGGAGAGATGATTACTTCTCTCTCTCTGACCGG	1125					

RESULT 3
US-10-195-071-1
; Sequence 1, Application US/10195071
; Publication No. US20030096271A1

GENERAL INFORMATION:
; APPLICANT: Origene Technologies
; TITLE OF INVENTION: Full-Length Serine Protein Kinase in Brain and Pancreas
; FILE REFERENCE: 16U 101 C1
; CURRENT APPLICATION NUMBER: US/10/195,071
; CURRENT FILING DATE: 2002-07-15
; PRIOR APPLICATION NUMBER: US 09/930,181
; PRIOR FILING DATE: 2001-08-16
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1
; LENGTH: 2908
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (106)..(2112)
; OTHER INFORMATION:
US-10-195-071-1

Query Match 99.9%; Score 2005.4; DB 14; Length 2908;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2006; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy	1	ATGACATCGACGGGGAAGGACGGCGCGCAGAGCA	CGCGCAGTATGTTGGGCCCTACCGG	60
Db	106	ATGACATCGACGGGGAAGGACGGCGCGCAGAGCA	CGCGCAGTATGTTGGGCCCTACCGG	165
Qy	61	CTGGAGAAGACGCTGGGCAAGGGGACACAGGCT	CTGCTGAAGCTGGGGGTTCACTGGCTC	120
Db	166	CTGGAGAAGACGCTGGGCAAGGGGACACAGGCT	CTGCTGAAGCTGGGGGTTCACTGGCTC	225
Qy	121	ACCTGCCAGAGGTGGCCATCAAGATCGTCAAC	CGCTGAGAAGCTCAGCGAGTGGTGCTG	180
Db	226	ACCTGCCAGAGGTGGCCATCAAGATCGTCAAC	CGCTGAGAAGCTCAGCGAGTGGTGCTG	285
Qy	181	ATGAAGTGGAGCGGGAGATCGGATCGTGAAG	CTCATTTAGACACCCCGCTCTTAAG	240
Db	286	ATGAAGTGGAGCGGGAGATCGGATCGTGAAG	CTCATTTAGACACCCCGCTCTTAAG	345
Qy	241	CTGCACGACGCTTTATGAACAACAAAAATAT	TGTACCTGGTGTAGAACACGTTGTCAGGT	300
Db	346	CTGCACGACGCTTTATGAACAACAAAAATAT	TGTACCTGGTGTAGAACACGTTGTCAGGT	405
Qy	301	GGTGAGCTCTTTCGACTACCTGGTGAAGAAG	GGGGAGGCTGACCGCTAAGAGGCTCGGAAG	360
Db	406	GGTGAGCTCTTTCGACTACCTGGTGAAGAAG	GGGGAGGCTGACCGCTAAGAGGCTCGGAAG	465
Qy	361	TTCTTCCGGGAGATCATCTCGCTGGACTTCT	CGCACAGCCACTCCATATGCCACAGG	420
Db	466	TTCTTCCGGGAGATCATCTCGCTGGACTTCT	CGCACAGCCACTCCATATGCCACAGG	525
Qy	421	GATCTGAACCTGAAACCTCTCTGTGACAGAG	AAGAACAACTCCGCTATCGCAGACTTT	480
Db	526	GATCTGAACCTGAAACCTCTCTGTGACAGAG	AAGAACAACTCCGCTATCGCAGACTTT	585
Qy	481	GGCATGCGGCTCCCTGCAAGTTGGCGACAG	CTGTGAGACACAGCTGTGGGTCCCCC	540
Db	586	GGCATGCGGCTCCCTGCAAGTTGGCGACAG	CTGTGAGACACAGCTGTGGGTCCCCC	645
Qy	541	TACGCTGCCCGGAGGTGATCCGGGGGAGAG	TATGACGGCGGAGGGCGGAGCTGTGG	600
Db	646	TACGCTGCCCGGAGGTGATCCGGGGGAGAG	TATGACGGCGGAGGGCGGAGCTGTGG	705
Qy	601	AGCTGCGGCGTCACTCTGCTGCGCTTGTCT	GCTGGGGGCTCTGCGCTTCGACGATGACAAC	660
Db	706	AGCTGCGGCGTCACTCTGCTGCGCTTGTCT	GCTGGGGGCTCTGCGCTTCGACGATGACAAC	765
Qy	661	TTGGGACAGCTGCTGGAGAAAGTGGAAG	CGGGCGGTGTTCCACATGCCGCACTTTATCCCG	720
Db	766	TTGGGACAGCTGCTGGAGAAAGTGGAAG	CGGGCGGTGTTCCACATGCCGCACTTTATCCCG	825
Qy	721	CCCGACTGCCAGAGTCTCTACGGGGCATG	ATCGAGGTGGACGGCGCGCCCTCAG	780

Db	826	CCCGACTGCGCAGAGTCTGCTACGGGGCATG	ATCGAGGTGGACGGCGCGCTCTCAG	885
Qy	781	CTAGAGCACAATTCAGAAACACATATGTA	TATAGGGGGCAAGAAATGAGCCGAAACAGAG	840
Db	886	CTAGAGCACAATTCAGAAACACATATGTA	TATAGGGGGCAAGAAATGAGCCGAAACAGAG	945
Qy	841	CAGCCCAATTCCTCGCAAGGTGAGATCCG	CTCGCTGCGCCAGCTGGAGACATCGACCCC	900
Db	946	CAGCCCAATTCCTCGCAAGGTGAGATCCG	CTCGCTGCGCCAGCTGGAGACATCGACCCC	1005
Qy	901	GACGTGCTGGACAGATGCACTCACTGGG	CTGCTTCCGAGACCGCAACAGCTGCTGAG	960
Db	1006	GACGTGCTGGACAGATGCACTCACTGGG	CTGCTTCCGAGACCGCAACAGCTGCTGAG	1065
Qy	961	GACCTGCTGCTCCGAGAGGAGAACACCA	GAGAGATGATTTACTTCTCTCTGACCCGG	1020
Db	1066	GACCTGCTGCTCCGAGAGGAGAACACCA	GAGAGATGATTTACTTCTCTCTGACCCGG	1125
Qy	1021	AAAGAAAGTACCCGAGCAGAGGATGAG	ACCTGCCCCCCCCCGGAACAGATAGACCT	1080
Db	1126	AAAGAAAGGTACCCGAGCAGAGGATGAG	ACCTGCCCCCCCCCGGAACAGATAGACCT	1185
Qy	1081	CCCGGAGAGCTGTGCACTCCCGGATGCT	GAACCGGCAAGCGCGGCGCAAGACGC	1140
Db	1186	CCCGGAGAGCTGTGCACTCCCGGATGCT	GAACCGGCAAGCGCGGCGCAAGACGC	1245
Qy	1141	AAATCATCGAGGTGCTCAGCGTGACG	ACGCGGCTCCCGGCTGCTCGCGGCGGCGC	1200
Db	1246	AAATCATCGAGGTGCTCAGCGTGACG	ACGCGGCTCCCGGCTGCTCGCGGCGGCGC	1305
Qy	1201	ATTGAGATGGCCAGCAGCGCCAGAGGT	CTCGGTCCATCAGCGGTGCTCTCAGGCTTT	1260
Db	1306	ATTGAGATGGCCAGCAGCGCCAGAGGT	CTCGGTCCATCAGCGGTGCTCTCAGGCTTT	1365
Qy	1261	TCACACAGCCCACTCAGCAGACCCCGG	GTGACCCCTCACCAAGGGGCGAGTCCC	1320
Db	1366	TCACACAGCCCACTCAGCAGACCCCGG	GTGACCCCTCACCAAGGGGCGAGTCCC	1425
Qy	1321	CTCCCCCAACCCCAAGGGGACACCTGT	CTCAACGCGCAAGAGGAGCGCGCTGACGCCC	1380
Db	1426	CTCCCCCAACCCCAAGGGGACACCTGT	CTCAACGCGCAAGAGGAGCGCGCTGACGCCC	1485
Qy	1381	AACCCACAGCCCGCTGCAAGCCAGCG	CTGAGGGGTGCTTGGAGGGCGGCTCAAC	1440
Db	1486	AACCCACAGCCCGCTGCAAGCCAGCG	CTGAGGGGTGCTTGGAGGGCGGCTCAAC	1545
Qy	1441	TCCATCAAGAAACAGCTTTCTGGGCT	CCACCCCGCTTCCACCGCGGAACTGCAAGTTCCG	1500
Db	1546	TCCATCAAGAAACAGCTTTCTGGGCT	CCACCCCGCTTCCACCGCGGAACTGCAAGTTCCG	1605
Qy	1501	ACCCCGGAGAGATGTCCAACCTTGAC	CACAGAGTCTGCTCCAGAGCTGGCGAAGAGTCC	1560
Db	1606	ACCCCGGAGAGATGTCCAACCTTGAC	CACAGAGTCTGCTCCAGAGCTGGCGAAGAGTCC	1665
Qy	1561	TGCTTTGGGAACTTCACTACGCTGGA	AGAGGAGCAGATCTTCGTGTCATCAAGAC	1620
Db	1666	TGCTTTGGGAACTTCACTACGCTGGA	AGAGGAGCAGATCTTCGTGTCATCAAGAC	1725
Qy	1621	AAACCTCTGAGCTCCATCAAGGCTGA	CATCGTGCAGCCCTTCTGTCGATTTCCAGTCTC	1680
Db	1726	AAACCTCTGAGCTCCATCAAGGCTGA	CATCGTGCAGCCCTTCTGTCGATTTCCAGTCTC	1785
Qy	1681	AGCCACAGCGTCTCTCCAAACAGAG	CTTCGGGCGGAGTACAGGCGCCACGGGGGGCCA	1740
Db	1786	AGCCACAGCGTCTCTCCAAACAGAG	CTTCGGGCGGAGTACAGGCGCCACGGGGGGCCA	1845
Qy	1741	GCGGTGTTCCAGAGCGGTCAGATTCC	AGGTTGATACCTACACGAGGGTGGGGAG	1800
Db	1846	GCGGTGTTCCAGAGCGGTCAGATTCC	AGGTTGATACCTACACGAGGGTGGGGAG	1905
Qy	1801	GCGCAGAGGAGAACGCGACTACTCCG	TACCTCAGCTTCCAGGCGCCAGCCGT	1860

1557	QY	GTCTCGTGTGGAACTTTCATCAGCCTGGAGAGAGGAGCAGATCTTCGTGTGCATCAA	1616
1789	Db	GTCTCGTGTGGAACTTTCATCAGCCTGGAGAGAGGAGCAGATCTTCGTGTGCATCAA	1848
1617	QY	AGACAAACCTCTCAGCTCCATCAAGGCTGACATCGGTGCAGCCCTTCTGTTCGATTCGCCAG	1676
1849	Db	AGACAAACCTCTGAGCTCCATCAAGGCTGACATCGGTGCAGCCCTTCTGTTCGATTCGCCAG	1908
1677	QY	TCTCAGCCACAGCGTCAATCTCCAAACGAGCTTCCGGGCCGAGTACAAGGCCACCGGGGG	1736
1909	Db	TCTCAGCCACAGCGTCAATCTCCAAACGAGCTTCCGGGCCGAGTACAAGGCCACCGGGGG	1968
1737	QY	GCCAGCGGTGTTCCAGNAGCGGCTCAAGTTCAGGTTGATACCACTACACGAGGGGTGG	1796
1969	Db	GCCAGCGGTGTTCCAGNAGCGGCTCAAGTTCAGGTTGATACCACTACACGAGGGGTGG	2028
1797	QY	GGAGGCGCAGAGGAGAACCGGCATCTACTCCGTCACCTTCACCCCTGCTCTCAGGCCCCACG	1856
2029	Db	GGAGGCGCAGAGGAGAACCGGCATCTACTCCGTCACCTTCACCCCTGCTCTCAGGCCCCACG	2088
1857	QY	CCGTGCGTTCAAGAGGGTGTGGAGACCATCCAGGCCACGCTGCTGAGCAACACGACCC	1916
2089	Db	CCGTGCGTTCAAGAGGGTGTGGAGACCATCCAGGCCACGCTGCTGAGCAACACGACCC	2148
1917	QY	GCCTGCGGCCACGACCTTGTGCAGACACCACTTAATCTGTATCGAAATGATACGGGGGGCT	1976
2149	Db	GCCTGCGGCCACGACCTTGTGCAGACACCACTTAATCTGTATCGAAATGATACGGGGGGCT	2208
1977	QY	TTCCAAATGTGGAATTATCCCGAAAAGTTAA	2007
2209	Db	TTCCAAATGTGGAATTATCCCGAAAAGTTAA	2239

```

RESULT 5
US-10-195-071-3
; Sequence 3, Application US/10195071
; Publication NO. US20030096271A1
; GENERAL INFORMATION:
; APPLICANT: OriGene Technologies
; TITLE OF INVENTION: Full-Length Serine Protein Kinase in Brain and Pancreas
; FILE REFERENCE: 16U 101 C1
; CURRENT APPLICATION NUMBER: US/10/195,071
; CURRENT FILING DATE: 2002-07-15
; PRIOR APPLICATION NUMBER: US 09/930,181
; PRIOR FILING DATE: 2001-08-16
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 3
; LENGTH: 3364
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (482)..(2239)
; OTHER INFORMATION:
US-10-195-071-3

```

	Query Match	99.3%	Score 1993;	DB 14;	Length 3364;
	Best Local Similarity	99.8%;	Pred. No. 0;		
	Matches 2007; Conservative	0;	Mismatches	0; Indels	4; Gaps
	1;				
Qy	1	ATGACATCGACGGGGAAGGACGGCGGGCGGCGACGACGCGCAGTATGTTGGGCCCTACCGG	60		
Db	229	ATGACATCGACGGGGAAGGACGGCGGGCGGCGACGACGCGCAGTATGTTGGGCCCTACCGG	288		
Qy	61	CTGGAGAAAGACGCTCTGGGCGAAGGGGCGACAGAGTCTGGTGAAGCTGGGGGTTCACTGGGCTC	120		
Db	289	CTGGAGAAAGACGCTGGGCGAAGGGGCGACAGAGTCTGGTGAAGCTGGGGGTTCACTGGGCTC	348		
Qy	121	ACTTGCAGAAAGTGGGCATCAAGATCGTCAACCGTGAGAAAGCTCAACGAGTCGGTCTGCTG	180		
Db	349	ACTTGCAGAAAGTGGGCATCAAGATCGTCAACCGTGAGAAAGCTCAACGAGTCGGTCTGCTG	408		

181	ATGAAGGTGGAGCGGGAGATCCGGATCCTGAAAGCTCATTTGAGCAGCCCCACGCTCTCTAAAG	240
Qy		
409	ATGAAGGTGGAGCGGGAGATCCGGATCCTGAAAGCTCATTTGAGCAGCCCCACGCTCTCTAAAG	468
Db		
241	CTGCAAGCGTTTATGAAAAAATAATTT- - -GTACCTGGTGTCTAGAACACGCTGTC	296
Qy		
469	CTGCAAGCGTTTATGAAAAAATAATTTGAGGTACCTGGTGTCTAGAACACGCTGTC	528
Db		
297	AGTGGTGAAGCTTTGCACTA CTTGGTGAAGAAGGGGAGCTGAGCGCTTAAAGGAGCTCG	356
Qy		
529	AGTGGTGAAGCTTTGCACTA CTTGGTGAAGAAGGGGAGGCTGAGCGCTTAAAGGAGCTCG	588
Db		
357	GAAGTTCTTCGGCAGATCATCTCTGCGCTGCACTCTTCCACAGCCACTCCCATATGCCA	416
Qy		
589	GAAGTTCTTCGGCAGATCATCTCTGCGCTGCACTCTTCCACAGCCACTCCCATATGCCA	648
Db		
417	CAGGGATCTGAACCTGTAAAACTCTCTGTGGAACGAGAAAGAAACAACATCCGATCGCAGA	476
Qy		
649	CAGGGATCTGAACCTGTAAAACTCTCTGTGGAACGAGAAAGAAACAACATCCGATCGCAGA	708
Db		
477	CTTTGGCATGGCGTCCCTGCAAGTTGGCGACAGCCTGTTTGGAGACAGCTGTGGGTCCCC	536
Qy		
709	CTTTGGCATGGCGTCCCTGCAAGTTGGCGACAGCCTGTTTGGAGACAGCTGTGGGTCCCC	768
Db		
537	CCACTACGCTGCCCGAGTGATCCGGGGGAGAGTATGACGGCGCGAAGCGCGACGTC	596
Qy		
769	CCACTACGCTGCCCGAGTGATCCGGGGGAGAGTATGACGGCGCGAAGCGCGACGTC	828
Db		
597	GTGGAAGCTGGCGGTCAATCTCTGTTGCGCTTGTGTGGGGGCTCTGCCCTTGCAGCATGA	656
Qy		
829	GTGGAAGCTGGCGGTCAATCTCTGTTGCGCTTGTGTGGGGGCTCTGCCCTTGCAGCATGA	888
Db		
657	CAACTTGGCAGCAGCTCTCGAAGAGTGAAGCGGGGCGTGTTCACATCGCGACATTTAT	716
Qy		
889	CAACTTGGCAGCAGCTCTCGAAGAGTGAAGCGGGGCGTGTTCACATCGCGACATTTAT	948
Db		
717	CCCGCCCGACTGCGACAGATCTGCTACGGGGCATGATCGAGGTGGACCGCCGACCGCGCCT	776
Qy		
949	CCCGCCCGACTGCGACAGATCTGCTACGGGGCATGATCGAGGTGGACCGCCGACCGCGCCT	1008
Db		
777	CAGCTTAGACACATTCAGAAAACACATATGTGTATATAGGGGGCAGAAATGAGCCCGAAC	836
Qy		
1009	CAGCTTAGACACATTCAGAAAACACATATGTGTATATAGGGGGCAGAAATGAGCCCGAAC	1068
Db		
837	AGAGCAGCCCATTCCTCGAAAGGTGCAGATCCGCTCGCTGCCCAGCGCTGGAGGACATCGA	896
Qy		
1069	AGAGCAGCCCATTCCTCGAAAGGTGCAGATCCGCTCGCTGCCCAGCGCTGGAGGACATCGA	1128
Db		
897	CCCGCAGCTGCTGGACAGCATGCACTCACTGGGCTGCTTCGAGACCGCAACAAAGCTGCT	956
Qy		
1129	CCCGCAGCTGCTGGACAGCATGCACTCACTGGGCTGCTTCGAGACCGCAACAAAGCTGCT	1188
Db		
957	GCAAGGACCTGCTGTCGAGGAGGAGAACAGAGAGAGATGATTTACTTCTCTCTCTGGA	1016
Qy		
1189	GCAAGGACCTGCTGTCGAGGAGGAGAACAGAGAGAGATGATTTACTTCTCTCTCTGGA	1248
Db		
1017	CCGGAAGAAAGGTACCCGAGCCAGGAGGATGAGGACCTGCCCCCCCGGAAACAGATAGA	1076
Qy		
1249	CCGGAAGAAAGGTACCCGAGCCAGGAGGATGAGGACCTGCCCCCCCGGAAACAGATAGA	1308
Db		
1077	CCCTCCCGGAAGCGTGTGACCTCCCGATGCTGTAAACCGGCACGGCAAGCGGCGGCCAGA	1136
Qy		
1309	CCCTCCCGGAAGCGTGTGACCTCCCGATGCTGTGTAAACCGGCACGGCAAGCGGCGGCCAGA	1368
Db		
1137	ACGCAAAATCCATGAGGTGCTCAGCGGTGACGGACGGCGGCTCCCGGGTGCCTCGCGGGCG	1196
Qy		
1369	ACGCAAAATCCATGAGGTGCTCAGCGGTGACGGACGGCGGCTCCCGGGTGCCTCGCGGGCG	1428
Db		
1197	GGCCATTGAGATGGCCGACGAGCGGATGCTCGGTTCATCAGCGGTGCTCTCTCAGG	1256
Qy		
1429	GGCCATTGAGATGGCCGACGAGCGGATGCTCGGTTCATCAGCGGTGCTCTCTCAGG	1488
Db		

Db 570 CCGAAGCGGACGTGTGGAGCTGGCGGCTATCTCTGTTGGCTGTGTGGTGGGGCTCT 629
Qy 642 GCCCTTCGACGATGACAACTTGGACAGCTGTGGAGAAAGGTGAGCGGGGCGCTGTCCA 701
Db 630 GCCCTTCGACGATGACAACTTGGACAGCTGTGGAGAAAGGTGAGCGGGGCGCTGTCCA 699
Qy 702 CATGCCGCACTTTATCCCGCCGACTGCCAGAGTCTCTACGGGGCATGATCGAGGTGGA 761
Db 690 CATGCCGCACTTTATCCCGCCGACTGCCAGAGTCTCTACGGGGCATGATCGAGGTGGA 749
Qy 762 CGCGCGACGCGGCTCAGCTAGAGCACTTCAGAAACACATATGGTATATAGGGGCAA 821
Db 750 CGCGCGACGCGGCTCAGCTAGAGCACTTCAGAAACACATATGGTATATAGGGGCAA 809
Qy 822 GAATGAGCCGGAACAGAGAGCCCACTCTCGCAAGGTGCAGATCCGCTCGCTGCCAG 881
Db 810 GAATGAGCCGGAACAGAGAGCCCACTCTCGCAAGGTGCAGATCCGCTCGCTGCCAG 869
Qy 882 CTTGGAGGACATCGACCCCGACGCTGTGGACAGCATCTCACTGGGCTGCTTCCGAGA 941
Db 870 CTTGGAGGACATCGACCCCGACGCTGTGGACAGCATCTCACTGGGCTGCTTCCGAGA 929
Qy 942 CCGCAACAGCTGTGACGAGCTGTCTGTGGAGGAGAGAACAGAGAGATGATTTA 1001
Db 930 CCGCAACAGCTGTGACGAGCTGTCTGTGGAGGAGAGAACAGAGAGATGATTTA 989
Qy 1002 CTTCTCTCTCTGACCGGAAAGAGTACCCGAGCCAGGAGTACGAGACCTGCCCCC 1061
Db 990 CTTCTCTCTCTGACCGGAAAGAGTACCCGAGCCAGGAGTACGAGACCTGCCCCC 1049
Qy 1062 CCGAAGCAGATAGACCTCTCCCGAAGCGTGTGGACTCTCCCGATGCTGAACCGGCA 1121
Db 1050 CCGAAGCAGATAGACCTCTCCCGAAGCGTGTGGACTCTCCCGATGCTGAACCGGCA 1109
Qy 1122 CAAGCGGGGCGCAACGCAATCATGGAGTGTCTCAGGTGACGAGCGGCTCCCC 1181
Db 1110 CAAGCGGGGCGCAACGCAATCATGGAGTGTCTCAGGTGACGAGCGGCTCCCC 1169
Qy 1182 GGTGCTCTGCGGGGCGCATTTAGATGGCCAGCAGCGGCGGAGGCTCGGTCCATCAG 1241
Db 1170 GGTGCTCTGCGGGGCGCATTTAGATGGCCAGCAGCGGCGGAGGCTCGGTCCATCAG 1229
Qy 1242 CGGTGCTCTCTCAGGCTTTTCCACAGCCACTCAGCAGCCCGCGGTGACCCCTCACCC 1301
Db 1230 CGGTGCTCTCTCAGGCTTTTCCACAGCCACTCAGCAGCCCGCGGTGACCCCTCACCC 1289
Qy 1302 CTCACCAAGGGGAGTCCCTTCCCAACCCCAAGGGGACACTCTCCACAGCCCAAGGA 1361
Db 1290 CTCACCAAGGGGAGTCCCTTCCCAACCCCAAGGGGACACTCTCTCCACAGCCCAAGGA 1349
Qy 1362 GAGCCCGGCTGGCACGCCCAACCCCAACCCCGCTCCAGCCCGGAGGGGTGCC 1421
Db 1350 GAGCCCGGCTGGCACGCCCAACCCCAACCCCGCTCCAGCCCGGAGGGGTGCC 1409
Qy 1422 CTGAGGGGCGGCTCAACTCATCAAGAACAGCTTTCTGGGCTCAACCCGCTTCCACCG 1481
Db 1410 CTGAGGGGCGGCTCAACTCATCAAGAACAGCTTTCTGGGCTCAACCCGCTTCCACCG 1469
Qy 1482 CCGGAAACTGCAAGTTCCGAGCGCGGAGGAGATGTCACCACTGACACAGAGTCTGCC 1541
Db 1470 CCGGAAACTGCAAGTTCCGAGCGCGGAGGAGATGTCACCACTGACACAGAGTCTGCC 1529
Qy 1542 AGAGCTGGCAGAGTCTCTGTTTGGAACTTTCATCAGCCCTGGAGAGGAGGAGCAGAT 1601
Db 1530 AGAGCTGGCAGAGTCTCTGTTTGGAACTTTCATCAGCCCTGGAGAGGAGGAGCAGAT 1589
Qy 1602 CTTCTGTGTCTCAAGAGCAAACTCTGAGCTTCATCAAGGCTGACATCTGCGAGCCCTT 1661
Db 1590 CTTCTGTGTCTCAAGAGCAAACTCTGAGCTTCATCAAGGCTGACATCTGCGAGCCCTT 1649
Qy 1662 CTTGTGATTTCCAGCTCTCAGCCAGCGTCTATCTCCCAAGAGCTTCCGGGCGGAGTA 1721
Db 1650 CTTGTGATTTCCAGCTCTCAGCCAGCGTCTATCTCCCAAGAGCTTCCGGGCGGAGTA 1709

Qy 1722 CAAGGCCAGCGGGGCGCAGCGCTGTTCAGAAAGCGGTCAAGTTCCAGGTTGATATCAC 1781
Db 1710 CAAGGCCAGCGGGGCGCAGCGCTGTTCAGAAAGCGGTCAAGTTCCAGGTTGATATCAC 1769
Qy 1782 CTACACGAGGGTGGGGAGCGCAGAAAGGAGAACCGCATCTACTCCGTCACTTCACTT 1841
Db 1770 CTACACGAGGGTGGGGAGCGCAGAAAGGAGAACCGCATCTACTCCGTCACTTCACTT 1829
Qy 1842 GCTCTCAGGCCCCAGCGCTCGCTTCAAGAGGTTGGAGACATCCAGGCCAGCTGCT 1901
Db 1830 GCTCTCAGGCCCCAGCGCTCGCTTCAAGAGGTTGGAGACATCCAGGCCAGCTGCT 1889
Qy 1902 GAGCACACAGCACCGCCCTCGCGCCAGCACCTTGTGAGACACCACTAACTGTATGAAAT 1961
Db 1890 GAGCACACAGCACCGCCCTCGCGCCAGCACCTTGTGAGACACCACTAACTGTATGAAAT 1949
Qy 1962 GATGACGGGGCGCTTTCCAAATGTGGAATATTCGCAAAAGTTAA 2007
Db 1950 GATGACGGGGCGCTTTCCAAATGTGGAATATTCGCAAAAGTTAA 1995

RESULT 7

US-10-283-247-1

; Sequence 1, Application US/10283247

; Publication No. US20030119037A1

; GENERAL INFORMATION:

; APPLICANT: NEELAM, Beena et al.

; TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC

; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES

; FILE REFERENCE: CL001304

; CURRENT APPLICATION NUMBER: US/10/283,247

; CURRENT FILING DATE: 2002-10-30

; NUMBER OF SEQ ID NOS: 10

; SOFTWARE: fastseq for Windows Version 4.0

; SEQ ID NO 1

; LENGTH: 2025

; TYPE: DNA

; ORGANISM: Homo sapiens

; US-10-283-247-1

Query Match 96.7%; Score 1941.2; DB 14; Length 2025;

Best Local Similarity 99.8%; Pred. No. 0;

Matches 1943; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 ATGACATCGACGGGAGGACGGCGCGCAGCAGCAGCGAGTATGTTGGGCGCTACCGG 60
Db 1 ATGACATCGACGGGAGGAGGACGGCGCGCAGCAGCAGCGAGTATGTTGGGCGCTACCGG 60
Qy 61 CTGGAGAAGACGCTGGGCAAGGGGCGAGACAGGCTCTGTTGAAGCTGGGGGTTCACTGCGTC 120
Db 61 CTGGAGAAGACGCTGGGCAAGGGGCGAGACAGGCTCTGTTGAAGCTGGGGGTTCACTGCGTC 120
Qy 121 ACCTGCCAGAAAGTGGCCATCAAGATGCTCAACCGTGAAGCTCAGCGAGTGGTCTG 180
Db 121 ACCTGCCAGAAAGTGGCCATCAAGATGCTCAACCGTGAAGCTCAGCGAGTGGTCTG 180
Qy 181 ATGAAGTGGAGCGGAGATCGCGATCTTGAAGCTCATTGAGCACCCCGCCTCTAAAG 240
Db 181 ATGAAGTGGAGCGGAGATCGCGATCTTGAAGCTCATTGAGCACCCCGCCTCTAAAG 240
Qy 241 CTGCACGACGCTTTATGAAAACAAAAATATTTGTACCTGGTGTAGAACAGTGTCAAGGT 300
Db 241 CTGCACGACGCTTTATGAAAACAAAAATATTTGTACCTGGTGTAGAACAGTGTCAAGGT 300
Qy 301 GGTGAGCTCTTCGACTACCTGGTGAAGAGGGAGGCTGACCGCTAAGAGGCTCGGAAG 360
Db 301 GGTGAGCTCTTCGACTACCTGGTGAAGAGGGAGGCTGACCGCTAAGAGGCTCGGAAG 360
Qy 361 TTCTTCCGGCAGATCATCTCTGGCTGGACTTCTGCCACAGCCACTTCATATGCCACAGG 420
Db 361 TTCTTCCGGCAGATCATCTCTGGCTGGACTTCTGCCACAGCCACTTCATATGCCACAGG 420


```
QY 421 GATCTGAAACCTGAAACCTCTCTCTGGAAGAGAAACAAATCCGATCGCAGACTTT 480
Db 421 GATCTGAAACCTGAAACCTCTCTCTGGAAGAGAAACAAATCCGATCGCAGACTTT 480
QY 481 GGATGCGCTCCCTGCGAGTTGGCGACAGCTGTGGAGACCAAGCTGTGGSTCCCCCAC 540
Db 481 GGATGCGCTCCCTGCGAGTTGGCGACAGCTGTGGAGACCAAGCTGTGGSTCCCCCAC 540
QY 541 TAGCCTGCCCGAGGTGATCCCGGGGGAGAGATGACCGCGGAAAGCGGAGCGTGTGG 600
Db 541 TAGCCTGCCCGAGGTGATCCCGGGGGAGAGATGACCGCGGAAAGCGGAGCGTGTGG 600
QY 601 AGCTGCGGCGTCACTCTGTTGCGCTTCTGCTGGTGGGGCTCTGCCCTTCGAGAGATGAAC 660
Db 601 AGCTGCGGCGTCACTCTGTTGCGCTTCTGCTGGTGGGGCTCTGCCCTTCGAGAGATGAAC 660
QY 661 TTGCGACAGCTGTGGAGAGAGTGAAGCGGGCGTGTTCACATGCGCGCACTTTATCCCG 720
Db 661 TTGCGACAGCTGTGGAGAGAGTGAAGCGGGCGTGTTCACATGCGCGCACTTTATCCCG 720
QY 721 CCGGACTGCCAGAGTCTGCTACCGGGGATGATCGAGGTGGAACCGCGACGCGCTCACG 780
Db 721 CCGGACTGCCAGAGTCTGCTACCGGGGATGATCGAGGTGGAACCGCGACGCGCTCACG 780
QY 781 CTAGAGCACATTGAGAAACACATATGGTATATAGGGGGCAAGATGAGCCCGAACACAGAG 840
Db 781 CTAGAGCACATTGAGAAACACATATGGTATATAGGGGGCAAGATGAGCCCGAACACAGAG 840
QY 841 CAGGCCATCTCTCGAAGGTGCAGATCGCTCGCTGCCAGCTTGGAGGACATCGACCCC 900
Db 841 CAGGCCATCTCTCGAAGGTGCAGATCGCTCGCTGCCAGCTTGGAGGACATCGACCCC 900
QY 901 GACTGTGGAACAGCATGCACTCACTGGCTGTCTCCGAGACCGCAACAGCTGCTGAG 960
Db 901 GACTGTGGAACAGCATGCACTCACTGGCTGTCTCCGAGACCGCAACAGCTGCTGAG 960
QY 961 GACTGTGTCGAGGAGGAGAAACAGAGNAGATGATTTACTTCTCTCTCTGGACCGG 1020
Db 961 GACTGTGTCGAGGAGGAGAAACAGAGNAGATGATTTACTTCTCTCTCTGGACCGG 1020
QY 1021 AAAGAAAGGTACCGGACGAGGATGAGGACTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1080
Db 1021 AAAGAAAGGTACCGGACGAGGATGAGGACTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1080
QY 1081 CCGCGAAGCGTGTGGAATCCCGGATGCTGAAACCGGACGCGCAAGCGGCGGCAAGACG 1140
Db 1081 CCGCGAAGCGTGTGGAATCCCGGATGCTGAAACCGGACGCGCAAGCGGCGGCAAGACG 1140
QY 1141 AAATCCATGGAGTGTCTCAGCGTACGGAACGCGGCTCCCGGCTGCTGCGCGGCGG 1200
Db 1141 AAATCCATGGAGTGTCTCAGCGTACGGAACGCGGCTCCCGGCTGCTGCGCGGCGG 1200
QY 1201 ATTGAGATGGCCAGCAGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1260
Db 1201 ATTGAGATGGCCAGCAGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1260
QY 1261 TCCACAGCCCACTCAGCAGCCCCCGGTGACCCCTTCAACAGGCGGCGAGTCCC 1320
Db 1261 TCCACAGCCCACTCAGCAGCCCCCGGTGACCCCTTCAACAGGCGGCGAGTCCC 1320
QY 1321 CTCCCGACCCCGAGGAGACACTGTCCACAGCGCAAGAGGAGCCCGGTGGACGCGCC 1380
Db 1321 CTCCCGACCCCGAGGAGACACTGTCCACAGCGCAAGAGGAGCCCGGTGGACGCGCC 1380
QY 1381 AACCCCAAGCGGCTGCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1440
Db 1381 AACCCCAAGCGGCTGCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1440
QY 1441 TCCATCAAGAACAGCTTTCTGGGCTCAACCGGCTTCCACCGCGGAAACTGCAAGTTCCG 1500
Db 1441 TCCATCAAGAACAGCTTTCTGGGCTCAACCGGCTTCCACCGCGGAAACTGCAAGTTCCG 1500
```

```
QY 1501 AGCCCGGAGAGATGTCCAACTGACACAGAGTCTCTCCCGAGAGCTGGCGAAGAGTCC 1560
Db 1501 AGCCCGGAGAGATGTCCAACTGACACAGAGTCTCTCCCGAGAGCTGGCGAAGAGTCC 1560
QY 1561 TGCTTTGGGAACTTCATCAGCCTTGGAGAGGAGGAGCAGATCTTCGTGTGTCATCAAGAC 1620
Db 1561 TGCTTTGGGAACTTCATCAGCCTTGGAGAGGAGGAGCAGATCTTCGTGTGTCATCAAGAC 1620
QY 1621 AAACCTCTGAGCTCCATCAAGGCTGATCGTGACCGCTTCTGTCGATTTCCAGTCTC 1680
Db 1621 AAACCTCTGAGCTCCATCAAGGCTGATCGTGACCGCTTCTGTCGATTTCCAGTCTC 1680
QY 1681 AGCCACAGCTCATCTCCCAAACGAGCTTCCGGGCGGAGTACAAGGCGCCACGGGGGGCCA 1740
Db 1681 AGCCACAGCTCATCTCCCAAACGAGCTTCCGGGCGGAGTACAAGGCGCCACGGGGGGCCA 1740
QY 1741 GCGTGTTCAGAAAGCGGTCAAGTTCAGAGTTGATATCACTACACGAGAGGCTGGGGAG 1800
Db 1741 GCGTGTTCAGAAAGCGGTCAAGTTCAGAGTTGATATCACTACACGAGAGGCTGGGGAG 1800
QY 1801 GCGCAGAGGAGAACCGGCTACTCTCCGTCACTTCACTTCACTTCACTTCACTTCACTTCACT 1860
Db 1801 GCGCAGAGGAGAACCGGCTACTCTCCGTCACTTCACTTCACTTCACTTCACTTCACTTCACT 1860
QY 1861 CGCTTCAAGAGGCTGTGGAGACCATTCAGGCGGAGCTGCTGAGGACACACGACCGGCT 1920
Db 1861 CGCTTCAAGAGGCTGTGGAGACCATTCAGGCGGAGCTGCTGAGGACACACGACCGGCT 1920
QY 1921 GCGGCCCGAGCACTTGTGAGACACAC 1946
Db 1921 GCGGCCCGAGCACTTGTGAGACACAC 1946
```

RESULT 8

```
US-09-842-582-3
; Sequence 3, Application US/09842582
; Patent No. US20020155570A1
; GENERAL INFORMATION:
; APPLICANT: Millennium Pharmaceuticals, Inc.
; TITLE OF INVENTION: 2246, NOVEL PROTEIN KINASE MOLECULES AND
; FILE REFERENCE: 38155-20054.00
; CURRENT APPLICATION NUMBER: US/09/842,582
; CURRENT FILING DATE: 2001-04-25
; PRIOR APPLICATION NUMBER: US 60/199,391
; PRIOR FILING DATE: 2000-04-25
; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 2025
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-842-582-3
```

Query Match 96.8%; Score 1939.6; DB 10; Length 2025;

Best Local Similarity 99.8%; Pred. No. 0;

Matches 1942; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

```
QY 1 ATGACATCGACGGGGAAGGACGCGCGGCGGAGACAGAGTCTGGTGAAGCTGGGGGTTCATCGGTC 60
Db 1 ATGACATCGACGGGGAAGGACGCGCGGCGGAGACAGAGTCTGGTGAAGCTGGGGGTTCATCGGTC 60
QY 61 CTGGAGAAGACGCTGGGGCAAGGGGCGAGACAGAGTCTGGTGAAGCTGGGGGTTCATCGGTC 120
Db 61 CTGGAGAAGACGCTGGGGCAAGGGGCGAGACAGAGTCTGGTGAAGCTGGGGGTTCATCGGTC 120
QY 121 ACCTGCGAGAGGTGGCCATCAAGATCGTCAACCGTGAAGAGCTACGAGTGGTCTG 180
Db 121 ACCTGCGAGAGGTGGCCATCAAGATCGTCAACCGTGAAGAGCTACGAGTGGTCTG 180
QY 181 ATGAAGTGGAGCGGAGAGATCGGATCCCTGAAGCTCATTTGAGCACCCCGGCTCTAAAG 240
Db 181 ATGAAGTGGAGCGGAGAGATCGGATCCCTGAAGCTCATTTGAGCACCCCGGCTCTAAAG 240
```

181	ATGAGGTGAGCGGGAGATCGCGATCCTGAAAGCTCATTTGAGACACCCACGCTCTAAAG	240
241	CTGACGACGTTTATGAAAAAATAATTTGTA	300
241	CTGACGACGTTTATGAAAAAATAATTTGTA	300
301	GGTGAGCTCTTCACTACCTGGTGGAGAGGGGAGCTGACGCTTAAGGAGGCTCGGAAG	360
301	GGTGAGCTCTTCACTACCTGGTGGAGAGGGGAGCTGACGCTTAAGGAGGCTCGGAAG	360
361	TTCTTCCGGCAGATCATCTCTGCGCTGGACTTCTGCGACAGCCACTCCATATGCCACAGG	420
361	TTCTTCCGGCAGATCATCTCTGCGCTGGACTTCTGCGACAGCCACTCCATATGCCACAGG	420
421	GATCTGAACCTGAAACACTTCTGCTGGACGAGAGAACAACTATCGCATCGCAGACTTTT	480
421	GATCTGAACCTGAAACACTTCTGCTGGACGAGAGAACAACTATCGCATCGCAGACTTTT	480
481	GGCATGGCGTCCCTGCAGGTGGCGACAGCCTGTGTGGAGACAGCTGTGGGTCCCCCCAC	540
481	GGCATGGCGTCCCTGCAGGTGGCGACAGCCTGTGTGGAGACAGCTGTGGGTCCCCCCAC	540
541	TAGCCTTGCCCGAGGTGATCCGGGGGGAGAGTATGACGGCCGGAAGCGGACGTGTGG	600
541	TAGCCTTGCCCGAGGTGATCCGGGGGGAGAGTATGACGGCCGGAAGCGGACGTGTGG	600
601	AGCTGGCGGCTCATCTGTTTGCCTTCTGCTGGGGCTCTGCGCTTCCAGCATGACAAAC	660
601	AGCTGGCGGCTCATCTGTTTGCCTTCTGCTGGGGCTCTGCGCTTCCAGCATGACAAAC	660
661	TTGCGACAGCTGCTGGAGAGGTGAAGCGGGGCGTGTTCACATGCGGCACTTTATCCCG	720
661	TTGCGACAGCTGCTGGAGAGGTGAAGCGGGGCGTGTTCACATGCGGCACTTTATCCCG	720
721	CCGACTTCCAGAGTCTGCTACGGGGCATGATCGAGGTGAGCGCGACCGCGCTTCAAG	780
721	CCGACTTCCAGAGTCTGCTACGGGGCATGATCGAGGTGAGCGCGACCGCGCTTCAAG	780
781	CTAGAGCACATTTAGAAACACATATGTTATAGGGGGCAGAGTATGAGCCCGCAACAGAG	840
781	CTAGAGCACATTTAGAAACACATATGTTATAGGGGGCAGAGTATGAGCCCGCAACAGAG	840
841	CAGCCCATCTCTGCAAGGTGAGATCCGCTGCTGCCAGCTCGGAGGACATCGACCCC	900
841	CAGCCCATCTCTGCAAGGTGAGATCCGCTGCTGCCAGCTCGGAGGACATCGACCCC	900
901	GACGTGCTGACAGCATGCACTCACTGGGCTGCTTCGAGACCGCAACAAAGCTGCTGCAG	960
901	GACGTGCTGACAGCATGCACTCACTGGGCTGCTTCGAGACCGCAACAAAGCTGCTGCAG	960
961	GACTGTGCTCGAGAGGAGAACCCAGGAGAGATGATTTACTTCTCTCTCTGAGACCGG	1020
961	GACTGTGCTCGAGAGGAGAACCCAGGAGAGATGATTTACTTCTCTCTCTGAGACCGG	1020
1021	AAAGAAAGGTACCCGAGCCAGGAGATGAGACTCTGCCCGGACAGAGATAGACCTT	1080
1021	AAAGAAAGGTACCCGAGCCAGGAGATGAGACTCTGCCCGGACAGAGATAGACCTT	1080
1081	CCCCGGAAGCGTGTGGACTCTCCCGATGCTGAAACCGGCACGCGACGCGCGCAGAACGC	1140
1081	CCCCGGAAGCGTGTGGACTCTCCCGATGCTGAAACCGGCACGCGACGCGCGCAGAACGC	1140
1141	AAATCCATGGAGGTGCTCAGCGTGAACGAGCGGCGTCTCCCGGTGCTGCGCGCGGGCC	1200
1141	AAATCCATGGAGGTGCTCAGCGTGAACGAGCGGCGTCTCCCGGTGCTGCGCGCGGGCC	1200
1201	ATTGAGATGGCCACGACGGCCAGAGGTCTCGGTCCATCAGCGGTGCTCTCTCAGGCGCTT	1260
1201	ATTGAGATGGCCACGACGGCCAGAGGTCTCGGTCCATCAGCGGTGCTCTCTCAGGCGCTT	1260
1261	TCCACGACCCACTCAGACGACCCCGGGTGACCCCTCACCCCTTCAACAGGGGAGTCCC	1320
1261	TCCACGACCCACTCAGACGACCCCGGGTGACCCCTCACCCCTTCAACAGGGGAGTCCC	1320

1321	Qy	CTCCCAACCCCAAGGGGACACTGTGTCCACAGCCAAAGAGAGACCGGCTGCGACGCCC	1338
1321	Db	CTCCCAACCCCAAGGGGACACTGTGTCCACAGCCAAAGAGAGACCGGCTGCGACGCCC	1380
1381	Qy	AAACCCCAACGCCCCCGTCAGCCCCCAGCGTCGAGGGGTGCCCTGGAGGGCGCGGCTCAAC	1440
1381	Db	AAACCCCAACGCCCCCGTCAGCCCCCAGCGTCGAGGGGTGCCCTGGAGGGCGCGGCTCAAC	1440
1441	Qy	TCCATCAAGAAACAGCTTTCTGGGCTCACCCCGCTTCACACGCGCGAAACTGCAAGTTCCG	1500
1441	Db	TCCATCAAGAAACAGCTTTCTGGGCTCACCCCGCTTCACACGCGCGAAACTGCAAGTTCCG	1500
1501	Qy	ACGCCGAGGAGATGTCCAAACCTGACACCAAGAGTGTCTCCAGACTTGGCGAAGAAAGTCC	1560
1501	Db	ACGCCGAGGAGATGTCCAAACCTGACACCAAGAGTGTCTCCAGACTTGGCGAAGAAAGTCC	1560
1561	Qy	TGGTTTGGGAACTTCATCAGCCTTGGAGAGGAGGAGCAGACTTCCTGTGGTTCATCAAGAGAC	1620
1561	Db	TGGTTTGGGAACTTCATCAGCCTTGGAGAGGAGGAGCAGACTTCCTGTGGTTCATCAAGAGAC	1620
1621	Qy	AAACCTCTGAGCTCCATCAAGGCTGACATCGTGGACGCGCTTCTGTGATTCCTCAGTCTC	1680
1621	Db	AAACCTCTGAGCTCCATCAAGGCTGACATCGTGGACGCGCTTCTGTGATTCCTCAGTCTC	1680
1681	Qy	AGCCACAGCGTCATCTCCAAAAGAGCTTCGGGCGCGAGTACAAGGCCACGGGGGGGCCA	1740
1681	Db	AGCCACAGCGTCATCTCCAAAAGAGCTTCGGGCGCGAGTACAAGGCCACGGGGGGGCCA	1740
1741	Qy	GCCTGTTCAGAAACCGGTCAAGTTTCCAGTTTGATATCACTACACGGAGGGTGGGGAG	1800
1741	Db	GCCTGTTCAGAAACCGGTCAAGTTTCCAGTTTGATATCACTACACGGAGGGTGGGGAG	1800
1801	Qy	GCGCAGAAGGAGAACCGGCATCTACTCCGTCACTTCACTCCGTCTCAGGCCCCAGCCGT	1860
1801	Db	GCGCAGAAGGAGAACCGGCATCTACTCCGTCACTTCACTCCGTCTCAGGCCCCAGCCGT	1860
1861	Qy	CGCTTCAAGAGGGTGTGGAGACCATCCAGGCCCGAGCTGTGAGCACACAGACCCGCT	1920
1861	Db	CGCTTCAAGAGGGTGTGGAGACCATCCAGGCCCGAGCTGTGAGCACACAGACCCGCT	1920
1921	Qy	CGGCCCCAGCACTTGTCCAGACACAC	1946
1921	Db	CGGCCCCAGCACTTGTCCAGACACCC	1946

```

RESULT 9
US-09-842-582-1
; Sequence 1, Application US/09842582
; Patent No. US20020155570A1
; GENERAL INFORMATION:
; APPLICANT: Millennium Pharmaceuticals, Inc.
; APPLICANT: Meyers, Rachel
; TITLE OF INVENTION: 2246, NOVEL PROTEIN KINASE MOLECULES AND
; TITLE OF INVENTION: USES THEREFOR
; FILE REFERENCE: 38155-20054.00
; CURRENT APPLICATION NUMBER: US/09/842,582
; CURRENT FILING DATE: 2001-04-25
; PRIOR APPLICATION NUMBER: US 60/199,391
; PRIOR FILING DATE: 2000-04-25
; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 2219
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (53) ... (2077)
US-09-842-582-1

```

Query Match 96.6%; Score 1939.6; DB 10; Length 2219;
Best Local Similarity 99.8%; Pred. No. 0;

Matches 1942; Conservative 0; Mismatches 4; Indels 0; Gaps 0;			
Qy	1	ATGACATCGACGGGAGGAGCGCGCGCGCAGACGCGAGTATGTTGGGCCCTACCGG	60
Db	53	ATGACATCGACGGGAGGAGCGCGCGCGCAGACGCGAGTATGTTGGGCCCTACCGG	112
Qy	61	CTGAGAGAAACGCTGGGCAAGGGGAGACAGGTCTGGTGAAGCTGGGGGTTCACTGGCTC	120
Db	113	CTGAGAGAAACGCTGGGCAAGGGGAGACAGGTCTGGTGAAGCTGGGGGTTCACTGGCTC	172
Qy	121	ACCTGCCAGAGAGTGGCCATCAAGATCTCAACCGTGAAGAGCTCAGCGAGTGGCTG	180
Db	173	ACCTGCCAGAGAGTGGCCATCAAGATCTCAACCGTGAAGAGCTCAGCGAGTGGCTG	232
Qy	181	ATGAAGGTGGAGCGGAGATCGCGATCTGAAGCTCATTTGAGCAACCCCAAGTCTTAAAG	240
Db	233	ATGAAGGTGGAGCGGAGATCGCGATCTGAAGCTCATTTGAGCAACCCCAAGTCTTAAAG	292
Qy	241	CTGACAGACGTTTATGAAGAAACAAAAATATTTGTACCTGGTGTGAGAACACGTTCAAGGT	300
Db	293	CTGACAGACGTTTATGAAGAAACAAAAATATTTGTACCTGGTGTGAGAACACGTTCAAGGT	352
Qy	301	GGTGAAGCTCTCGACTACCTGGTGAAGAGGGAGGCTGACGCTTAAGAGGCTCGGAG	360
Db	353	GGTGAAGCTCTCGACTACCTGGTGAAGAGGGAGGCTGACGCTTAAGAGGCTCGGAG	412
Qy	361	TTCTTTCCGCGAGATCATCTCTGCGCTGGACTTCTGCCACAGCACTCATATGCCACAGG	420
Db	413	TTCTTTCCGCGAGATCATCTCTGCGCTGGACTTCTGCCACAGCACTCATATGCCACAGG	472
Qy	421	GATCTGAACCTTGAAACCTCTCTGTGACGAGAGAAACATCCGCGATCGCAGACTTT	480
Db	473	GATCTGAACCTTGAAACCTCTCTGTGACGAGAGAAACATCCGCGATCGCAGACTTT	532
Qy	481	GGCATGGCGTCCCTCGAGGTTGGCAGACGCTGTTGGAGACAGCTGGGTCCTCCCGAC	540
Db	533	GGCATGGCGTCCCTCGAGGTTGGCAGACGCTGTTGGAGACAGCTGGGTCCTCCCGAC	592
Qy	541	TACGCTCCCGGAGGTGATCCGGGGGAGAGTATGACGGCGGAGAGCGGAGCTGTGG	600
Db	593	TACGCTCCCGGAGGTGATCCGGGGGAGAGTATGACGGCGGAGAGCGGAGCTGTGG	652
Qy	601	AGCTCGGGGTCATCTGTTGCCCTTGTGTTGGGGGCTCTGCCCTTCGACGATGACAAAC	660
Db	653	AGCTCGGGGTCATCTGTTGCCCTTGTGTTGGGGGCTCTGCCCTTCGACGATGACAAAC	712
Qy	661	TTGCGACAGCTGCTGGAGAGGTGAAGCGGGCGTGTTCACATCGCGCATTTATCCCG	720
Db	713	TTGCGACAGCTGCTGGAGAGGTGAAGCGGGCGTGTTCACATCGCGCATTTATCCCG	772
Qy	721	CCCGACTGCGAGGCTGCTGACGGGGCATGATCGAGGTGGAGCGCGCACGCGCCTCACG	780
Db	773	CCCGACTGCGAGGCTGCTGACGGGGCATGAGCGAGGTGGAGCGCGCACGCGCCTCACG	832
Qy	781	CTPAGAGCAATTCAGAAACACATATGGTATATAGGGGGCAAGAAATGAGCCGGAACAGAG	840
Db	833	CTPAGAGCAATTCAGAAACACATATGGTATATAGGGGGCAAGAAATGAGCCGGAACAGAG	892
Qy	841	CAGCCCATTCCTCGCAAGGTGAGATCCGCTCGCTGCCCGCCAGCTGGAGCATCGACCCC	900
Db	893	CAGCCCATTCCTCGCAAGGTGAGATCCGCTCGCTGCCCGCCAGCTGGAGCATCGACCCC	952
Qy	901	GACGTGCTGGACAGATGCACTCACTGGGCTGCTTCCGAGAGCCGAAACAGCTGCTGAG	960
Db	953	GACGTGCTGGACAGATGCACTCACTGGGCTGCTTCCGAGAGCCGAAACAGCTGCTGAG	1012
Qy	961	GACCTGCTGCTCGAGAGGAGAACACAGGAGAAAGATGATTTACTTCTCTCTGGACCGG	1020
Db	1013	GACCTGCTGCTCGAGAGGAGAACACAGGAGAAAGATGATTTACTTCTCTCTGGACCGG	1072
Qy	1021	AAAGAAAGTACCCGAGCCAGGAGATGAGGACTGACCTGCCCGCCCGGAGACGATACCT	1080
Db	1073	AAAGAAAGTACCCGAGCCAGGAGATGAGGACTGACCTGCCCGCCCGGAGACGATACCT	1132

Qy	1081	CCCCGAAGCGTGTGACTCCCGGATGCTGAACCGGCACGCGCAAGCGGCGCCCGAGAACGC	1140
Db	1133	CCCCGAAGCGTGTGACTCCCGGATGCTGAACCGGCACGCGCAAGCGGCGCCCGAGAACGC	1192
Qy	1141	AAATCCATGAGAGTGTCTCAGCGTGAAGGACGCGGGTCCCGGTGCTGCGCGCGCGGCC	1200
Db	1193	AAATCCATGAGAGTGTCTCAGCGTGAAGGACGCGGGTCCCGGTGCTGCGCGCGCGGCC	1252
Qy	1201	ATTGAGATGSCCCACGACGSCCAGAGGTCTCGGTCCATCAGCGGTGCTCTCTCAGGCTT	1260
Db	1253	ATTGAGATGSCCCACGACGSCCAGAGGTCTCGGTCCATCAGCGGTGCTCTCTCAGGCTT	1312
Qy	1261	TCCACCAAGCCCACTCAGCAGCCCCCGGGTGAACCTTCAACCCCTCACCAAGGGGAGTCCC	1320
Db	1313	TCCACCAAGCCCACTCAGCAGCCCCCGGGTGAACCTTCAACCCCTCACCAAGGGGAGTCCC	1372
Qy	1321	CTCCCAACCCCAAGGGGACACCTGTCCACACGCGCAAGGAGAGCCCGGTGGCAGCGCC	1380
Db	1373	CTCCCAACCCCAAGGGGACACCTGTCCACACGCGCAAGGAGAGCCCGGTGGCAGCGCC	1432
Qy	1381	AACCCCAAGCCCGCTCCAGCCCGGCTCGAGGGGTGCCCTGGAGGGGCGGCTCAAC	1440
Db	1433	AACCCCAAGCCCGCTCCAGCCCGGCTCGAGGGGTGCCCTGGAGGGGCGGCTCAAC	1492
Qy	1441	TCCATCAAGAACAGCTTTCTGGGCTCACCCCGCTTCCACCGCGGAAACTGCAAGTTCCG	1500
Db	1493	TCCATCAAGAACAGCTTTCTGGGCTCACCCCGCTTCCACCGCGGAAACTGCAAGTTCCG	1552
Qy	1501	ACGCGGAGAGATGTCCAAACCTGACACAGAGTGTGTCCTCCAGAGCTGGCGAAGAGTCC	1560
Db	1553	ACGCGGAGAGATGTCCAAACCTGACACAGAGTGTGTCCTCCAGAGCTGGCGAAGAGTCC	1612
Qy	1561	TGGTTTGGAACTTCATCAGCTTGGAGAGGAGGAGAGATCTTCGTGTCATCAAGAGAC	1620
Db	1613	TGGTTTGGAACTTCATCAGCTTGGAGAGGAGGAGAGATCTTCGTGTCATCAAGAGAC	1672
Qy	1621	AAACCTCTGAGCTCCATCAAGGCTGACATGTCACGCTTCTGTCGATTTCCAGTCTC	1680
Db	1673	AAACCTCTGAGCTCCATCAAGGCTGACATGTCACGCTTCTGTCGATTTCCAGTCTC	1732
Qy	1681	AGCAGAGCTCATCTCCCAACGAGCTTCCGGCGCGAGTACAGGCCACGCGGGGGGCCA	1740
Db	1733	AGCAGAGCTCATCTCCCAACGAGCTTCCGGCGCGAGTACAGGCCACGCGGGGGGCCA	1792
Qy	1741	GCCGTGTTCCAGAGCGGTCAGTTCAGGTTGATATCACTACAGAGGTTGGGAG	1800
Db	1793	GCCGTGTTCCAGAGCGGTCAGTTCAGGTTGATATCACTACAGAGGTTGGGAG	1852
Qy	1801	GCGCAGAGGAGAACGGCATCTACTCCGTCAACCTTCAACCTGCTCTCAGGCCCCAGCGGT	1860
Db	1853	GCGCAGAGGAGAACGGCATCTACTCCGTCAACCTTCAACCTGCTCTCAGGCCCCAGCGGT	1912
Qy	1861	CGCTTCAAGAGGTTGGAGACCATCCAGGCCAGCTGTGTAGACACACAGACCCCGCT	1920
Db	1913	CGCTTCAAGAGGTTGGAGACCATCCAGGCCAGCTGTGTAGACACACAGACCCCGCT	1972
Qy	1921	GCGSCCCAGCACTTGTTCAGACACCCAC	1946
Db	1973	GCGSCCCAGCACTTGTTCAGAACCCCC	1998

RESULT 10
US-10-054-579-3
; Sequence 3, Application US/10054579
; Publication No. US20020137913A1
; GENERAL INFORMATION:
; APPLICANT: Turner, C. Alexander Jr.
; APPLICANT: Mathur, Brian
; TITLE OF INVENTION: No. US20020137913A1el Human Kinases and Polynucleotides Encoding t
; FILE REFERENCE: LEX-0300-USA
; CURRENT APPLICATION NUMBER: US/10/054, 579
; CURRENT FILING DATE: 2002-01-22

; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES
; FILE OF INVENTION: THERBOSF
; FILE REFERENCE: CL001304
; CURRENT APPLICATION NUMBER: US/10/283,247
; CURRENT FILING DATE: 2002-10-30
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 1911
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-283-247-4

Query Match 84.0%; Score 1685.2; DB 14; Length 1911;
Best Local Similarity 96.2%; Pred. No. 0;
Matches 1763; Conservative 0; Mismatches 3; Indels 66; Gaps 1;

Qy	181	ATGAAGGTGGAGCGGAGATCCGATCTCTGAAGCTCATTTGAGCACCCTCAAGTCTCTAAAG	240
Db	1	ATGAAGGTGGAGCGGAGATCCGATCTCTGAAGCTCATTTGAGCACCCTCAAGTCTCTAAAG	60
Qy	241	CTGCACGAGCTTTATGAACAAAATAATTTGTACCTGGTGTAGAACAGTGTCAAGT	300
Db	61	CTGCACGAGCTTTATGAACAAAATAATTTGTACCTGGTGTAGAACAGTGTCAAGT	120
Qy	301	GGTGAGCTCTTCGACTACCTGGTGAAGAGGGAGGCTGACGCCCTAAGAGGCTCGGAAG	360
Db	121	GGTGAGCTCTTCGACTACCTGGTGAAGAGGGAGGCTGACGCCCTAAGAGGCTCGGAAG	180
Qy	361	TTCTCCCGCAGATCATCTCTCGCTGTGACTTTCTGCCACAGCACTCCATATGCCACAGG	420
Db	181	TTCTCCCGCAGATCATCTCTCGCTGTGACTTTCTGCCACAGCACTCCATATGCCACAGG	240
Qy	421	GATCTGAACCTGAAACCTCTCTGTCGACGAGAGAGCAACATCCGATCGCAGACTTT	480
Db	241	GATCTGAACCTGAAACCTCTCTGTCGACGAGAGAGCAACATCCGATCGCAGACTTT	300
Qy	481	GGCATGGCTCTCCCTGAGGTTGGCGACAGCTGTTGGAGACAGCTGTGGGTCCCCCAC	540
Db	301	GGCATGGCTCTCCCTGAGGTTGGCGACAGCTGTTGGAGACAGCTGTGGGTCCCCCAC	360
Qy	541	TACGCTGCCCCGAGTGATTCGGGGGAGAGATGACCGCGGAGAGCGGACGCTGTGG	600
Db	361	TACGCTGCCCCGAGTGATTCGGGGGAGAGATGACCGCGGAGAGCGGACGCTGTGG	420
Qy	601	AGCTCGGGGTCACTCTGCTTCCCTTGTCTGCTGGGGGCTCTGCCCTTCGACGATCAAC	660
Db	421	AGCTCGGGGTCACTCTGCTTCCCTTGTCTGCTGGGGGCTCTGCCCTTCGACGATCAAC	480
Qy	661	TTTGACAGCTCTGAGAGAGGTGAAGCGGGCGTGTTCACATGCCCGCACTTTATCCCG	720
Db	481	TTTGACAGCTCTGAGAGAGGTGAAGCGGGCGTGTTCACATGCCCGCACTTTATCCCG	540
Qy	721	CCCGACGTCAGAGTCTGTCAGGGGATGATCGAGGTGAGACCGCGCTCAGCGCTCAG	780
Db	541	CCCGACGTCAGAGTCTGTCAGGGGATGATCGAGGTGAGACCGCGCTCAGCGCTCAG	600
Qy	781	CTAGAGCATTTCAGAAACACATATGTTATAGGGGCGAAGATGAGCCCGAACAGAG	840
Db	601	CTAGAGCATTTCAGAAACACATATGTTATAGGGGCGAAGATGAGCCCGAACAGAG	660
Qy	841	CAGCCCATTCCTCGCAAGGTGCAGATCCGCTCGCTGCCAGCTCGAGGAGCATCGACCCC	900
Db	661	CAGCCCATTCCTCGCAAGGTGCAGATCCGCTCGCTGCCAGCTCGAGGAGCATCGACCCC	720
Qy	901	GACGTGTGAGCAGATGACATCTGAGGCTCTTCGAGACCGCAACAGCTGTGAG	960
Db	721	GACGTGTGAGCAGATGACATCTGAGGCTCTTCGAGACCGCAACAGCTGTGAG	780
Qy	961	GACCTGTCTCGAGGAGAGAACAGGAGAGATGATTTACTTCTCTCTCTGACCGG	1020
Db	781	GACCTGTCTCGAGGAGAGAACAGGAGAGATGATTTACTTCTCTCTCTGACCGG	840

RESULT 12

US-10-116-326-1
; Sequence 1, Application US/10116326
; Publication No. US2003016689A1

Qy	1021	AAAGAAAGGTACCGAGCCAGGAGGATGAGGACTGCTGCCCCCGGAAACGAGATAGACCT	1080
Db	841	AAAGAAAGGTACCGAGCCAGGAGGATGAGGACTGCTGCCCCCGGAAACGAGATAGACCT	900
Qy	1081	CCCGGAAGCGTGTGAGACTCCCGATGCTGAAACCGGCAACGGCAAGCGCGGCCAGAACGC	1140
Db	901	CCCGGAAGCGTGTGAGACTCCCGATGCTGAAACCGGCAACGGCAAGCGCGGCCAGAACGC	960
Qy	1141	AAATCCATGAGGTGCTCAGCGTACCGAGCGGGCTCCCGGTGCTTGGCGCGCGGGCC	1200
Db	961	AAATCCATGAGGTGCTCAGCGTACCGAGCGGGCTCCCGGTGCTTGGCGCGCGGGCC	1020
Qy	1201	ATTGAGATGGCCACGACGCGCAG-	1224
Db	1021	ATTGAGATGGCCACGACGCGCAGGATTAAGCAATGTTTCAAGTAAAGCCTTGGATATCGCT	1080
Qy	1225	-----AGTCTCGGTTCATCAGCGGTGCTCTCTCA	1254
Db	1081	GAGGCCCATCCCAATTCAGCAAGAAAGACAGGCTCTCGGTCCATCAGCGGTGCTCTCTCA	1140
Qy	1255	GGCTTTTCCACCCAGCCCACTCAGCAGCCCGCGGTGACCTCACCCTCACCAGGGGC	1314
Db	1141	GGCTTTTCCACCCAGCCCACTCAGCAGCCCGCGGTGACCTCACCCTCACCAGGGGC	1200
Qy	1315	AGTCCCTTCCCAACCCCAAGGGGACACCTGTCCACACGCGCAAGAGAGAGCCCGCTGGC	1374
Db	1201	AGTCCCTTCCCAACCCCAAGGGGACACCTGTCCACACGCGCAAGAGAGAGCCCGCTGGC	1260
Qy	1375	ACGCCCAACCCAGCCCGCTCAGCCCGCGGTGAGGGGTGCTTGGAGGGCGCGG	1434
Db	1261	ACGCCCAACCCAGCCCGCTCAGCCCGCGGTGAGGGGTGCTTGGAGGGCGCGG	1320
Qy	1435	CTCAACTCCATCAAGAAACAGCTTTCTGGGTTCACCCCGCTTCCACCGCGGAAACTGCA	1494
Db	1321	CTCAACTCCATCAAGAAACAGCTTTCTGGGTTCACCCCGCTTCCACCGCGGAAACTGCA	1380
Qy	1495	GTTCCGACGCGGAGGAGATGTCCAACTGACACAGAGTGTCTCCAGAGCTGGCGAAG	1554
Db	1381	GTTCCGACGCGGAGGAGATGTCCAACTGACACAGAGTGTCTCCAGAGCTGGCGAAG	1440
Qy	1555	AGTCTCTGTTTGGAACTTCATCAGCTGAGAGAGGAGGAGAGATCTTCTGTTGATC	1614
Db	1441	AGTCTCTGTTTGGAACTTCATCAGCTGAGAGAGGAGGAGAGATCTTCTGTTGATC	1500
Qy	1615	AAAGCAAAACCTCTGAGCTCCATCAAGCGTGACATCGTGACAGCTTCTCTGATTTCCC	1674
Db	1501	AAAGCAAAACCTCTGAGCTCCATCAAGCGTGACATCGTGACAGCTTCTCTGATTTCCC	1560
Qy	1675	AGTCTCAGCCACAGCGTCACTCTCCAAACGAGCTTCCGGGCGGAGTACAGGCCACCGGG	1734
Db	1561	AGTCTCAGCCACAGCGTCACTCTCCAAACGAGCTTCCGGGCGGAGTACAGGCCACCGGG	1620
Qy	1735	GGGCGAGCGGTTCAGAGAGCGGTCAAGTTCAGGTTGATATCATCTACACGAGGGT	1794
Db	1621	GGGCGAGCGGTTCAGAGAGCGGTCAAGTTCAGGTTGATATCATCTACACGAGGGT	1680
Qy	1795	GGGAGGCGCAGAGGAGAAAGCGCATCTCTCGTCACTTCCCTGCTCTCAGGGCCC	1854
Db	1681	GGGAGGCGCAGAGGAGAAAGCGCATCTCTCGTCACTTCCCTGCTCTCAGGGCCC	1740
Qy	1855	AGCGTCTGCTTCAAGAGGGTGTGAGACCATCAGGCCCGAGCTGTGTGAGCACACGAC	1914
Db	1741	AGCGTCTGCTTCAAGAGGGTGTGAGACCATCAGGCCCGAGCTGTGTGAGCACACGAC	1800
Qy	1915	CCGCTCGGGCCAGCACTTGTTCAGACACAC	1946
Db	1801	CCGCTCGGGCCAGCACTTGTTCAGAAACCCCC	1832

GENERAL INFORMATION:
; APPLICANT: Turner, C. Alexander Jr.
; APPLICANT: Mathur, Brian
; APPLICANT: Friddle, Carl Johan
; TITLE OF INVENTION: No. US2003016689A1el Human Kinases and Polynucleotides Encoding
; FILE REFERENCE: LEX-0332-USA
; CURRENT APPLICATION NUMBER: US/10/116,326
; PRIOR FILING DATE: 2002-04-04
; PRIOR APPLICATION NUMBER: US 60/282,036
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 2337
; TYPE: DNA
; ORGANISM: homo sapiens
US-10-116-326-1

Query Match 48.1%; Score 965; DB 12; Length 2337;
Best Local Similarity 70.0%; Pred. No. 3.1e-251;
Matches 1484; Conservative 0; Mismatches 440; Indels 195; Gaps 6;
Qy 29 CGAGCAGCCAGTATGTTGGGCTTACCGCTGAGAGACCTGGGCGAGGGGAGA 88
Db 74 CCCAGCACGCCCAATATGTTGGGCTTATCGGCTGGAAGACGCTGGGCAAGGAGAGA 133
Qy 89 CAGGCTGTTGAAGTGGGGGTTCACTGCGTCACTGCCAGAGGTGGCCATCAAGATCG 148
Db 134 CAGGCTGTTGAAGTGGGGTTCACCTGCGTCACTGCCAGAGGTGGCCATCAAGATCG 193
Qy 149 TCAACCGGTGAGAAGCTCAGCAGTGGTGTGATGAAGGTGGAGCGGAGATCCGATCC 208
Db 194 TGAACCGGGAGAGCTGTCCGAGTGGTGTGATGAAGGTGGAGCGGAGATCCGATCC 253
Qy 209 TGAAGCTATTGAGCACCACCCAGCTCTAAGCTGACAGCTTTATGAACAAATAAT 268
Db 254 TGAAGCTATTGAGCACCACCCAGCTCTAAGCTGACAGCTTTATGAACAAATAAT 313
Qy 269 ATTTGTACTGTGTAGAACACGTGTGAGTGTGAGCTCTTGCATCTACCTGTGAGA 328
Db 314 ATTTGTACTGTGTGTGAGCAGCTCTCGGGGGTGTGAGCTATTGCACTTCTGTGAAGA 373
Qy 329 AGGGAGGCTGACGCTTAAGAGGCTCGGAAGTTCTCCGCAAGATCATCTGTGCTGG 388
Db 374 AGGGAGGCTGACGCTTAAGAGGCTCGGAAGTTCTCCGCAAGATCATCTGTGCTGG 433
Qy 389 ACTTGTGCGCAGCAGCTTCCATATGCGCAGGATCTGAACCTGAAACCTCTGCTGG 448
Db 434 ACTTGTGCGCAGCAGCTTCCATATGCGCAGGATCTGAACCTGAAACCTCTGCTGG 493
Qy 449 ACGAAGAAACAACTCCGCAATCGCAGACTTTGGCATGGCGCTCCCTGAGGTGGCGACA 508
Db 494 ATGAGAAAACAACTCCGCAATCGCAGTTGCGACTTGGCATGGCGCTCCCTGAGGTGGCGACA 553
Qy 509 GCCTTTGGAGACAGCTGTGGTTCCTCCCACTACGCTTGGCCCGAGGTGATCCGGGGG 568
Db 554 GCCTTTGGAGACAGCTGTGGTTCCTCCCACTATGCGTGTCCAGAGGTGATTAAGGGG 613
Qy 569 AGAAGTATGACGCGGAGGCGGAGCTGTGGAGCTGGCGGTTCATCTGTTCGCTTGC 628
Db 614 AAAAATATGATGGCGCGCGGAGACATGTGGAGCTGTGGAGCTCATCTTTCGCGCTGC 673
Qy 629 TGGTGGGGCTCTGCGCTTCGACGATGACAACTTGGCAGCTGTGGAGAGGTGAAGC 688
Db 674 TCGTGGGGCTCTGCGCTTTGATGACGACAACTCCGCGAGCTGTGGAGAGGTGAAC 733
Qy 689 GGGGCGTGTTCACATGCGCACTTTATCCCGCCGAGCTGCGCAGAGTCTGTACGGGCA 748
Db 734 GGGGCGTGTTCACATGCGCACTTTATCCCGCCGAGTGTGGAGAGTGTGGAGAGT 793
Qy 749 TGATCGAGTGTGACGCGCGCTCCAGCTGACGATGAGCAATTCAGAAACACATATGT 808
Db 794 TGATCGAGTGTGACGCGCGCTCCAGCTGAGTGTGAGCAATTCAGAAACATCTTGT 853

Qy 809 ATATAGGGGGCAAGATGAGTGGCCGAAACC-----AGAGCAGCCCATTCCTCGAAGGTGC 862
Db 854 ACTAGGCGGGAAAACACGAGCCAGAGCCCTGCTGGAGCAGCCCTGCTGCCCGCGGTAG 913
Qy 863 AGATCGCTGCTGCCAGCCTGGAGGACATCGACCCCGAGCTGTGTGAGACAGCATGCAT 922
Db 914 CCATGCGAGCCTGCTCATCAACGGAGAGCTGGACCCCGACGCTCTTAGAGAGCATGGCAT 973
Qy 923 CACTGGGCTGCTTCCGAGACCGCAACAGCTGCTGCAGGACCTGCTGCTCCGAGGAGAGA 982
Db 974 CACTGGGCTGCTTCCGAGACCGGAGGCTGCTGCTGCGAGCTGCGCAGTGGAGAGAGA 1033
Qy 983 ACCAGGAGAGATGATTATCTTCTCTCTCTGACCGGAAAGAGGTATCCCGAGCAGG 1042
Db 1034 ACCAAGAAAGATGATATATTCTGCTTTTGGATCGGAAGGAGCGGTATCCAGCTGTG 1093
Qy 1043 AGATGAGACCTGCCCCCGGAAACGAGATAGACCTCCCGGAAAGCTGTGGATCC 1102
Db 1094 AGGACGAGACCTGCTCCCGGAATGATGTTGACCCCGGAAAGCTGTGGATTTCTC 1153
Qy 1103 CGATGCTGAACCGGACGCGCAAGCGGCGCAGAACGCAATCCATGGAGGTCTCAGCG 1162
Db 1154 CCATGCTGAGCCCTCAGGGAAGCGGACGAGCGGAAGTCCATGGAAGTCTTAGCA 1213
Qy 1163 TGAC-----GGACGGCGCTCCCGGTGCTTGGCGGGCGGCAATTGAGATGGCCC 1213
Db 1214 TCACCGATGCGGGGGTGGTGGCTCCCTGTACCCACCGGAGCGGCTTGGAGATGGCCC 1273
Qy 1214 AGCAGGCGAGAGTCTGGTCCATCAGCGGTGCTCTCAGGCTTTTCCAGCAGCCAC 1273
Db 1274 AGCAGGCGAGAGATCCGCTAGCTCAGTGGAGCTTCCAGCGGTCTGTCTCCAGCGCTC 1333
Qy 1274 TCAGCAGCCCGCGGT-----ACCCCAAGGGGACACCTGTCCACACGCA 1357
Db 1334 TAAGCAGCCCAAGAGTCCGCTCTTTCTTTTTCACCGGAGCGGGGGTGGAGATGAGG 1393
Qy 1290 ----- 1289
Db 1394 CTCGAGCGGGGCTCCCGACTTCCAAACGACAGCGTCTTCTCGGGGCGCCAGGG 1453
Qy 1290 -----GACCCCTCACCCCTTCCAAAGGGGAGTCCCTCCCG----- 1326
Db 1454 GTGGGGCGCGGGAGACGACCCCGCCCGCTCCAGTCCCTCCACACCCCTGCGCGGC 1513
Qy 1327 -----ACCCCAAGGGGACACCTGTCCACACGCA 1357
Db 1514 CCCAGGCTCCCGCGCTCTCTGCGGGGACCCCTTGTGACTCGCCTCTGCAAGGCCCC 1573
Qy 1358 AGGAGAGCCCGGTGGCAAGCCCAAGCCCGCGCTCCAGCCCC-----AGCGTCG 1411
Db 1574 GGGCCAGTCCACCGGACCCCGGGGACAAACACACCCCGCGCGCGGTGGCGTCG 1633
Qy 1412 GAGGGTGGCTGGAGGCGGGCTCAACTCCTCAAGAACAGCTTCTGGGCTCACCCC 1471
Db 1634 GGGGAGCGCTGGAGGAGTCTCAACTCCTCAAGAACAGCTTCTGGGCTCCCTC 1693
Qy 1472 GCTTCCACCCCGGAAACTGCAAGTTCGACGCGCGGAGGAGATGTCCAACTGACACAG 1531
Db 1694 GCTTTCACCGGCGCAGATGAGTCCCTACCGCTGAGGAGATGTCCAGCTTGAAGCCAG 1753
Qy 1532 AGTGTCTCCAGAGCTGGCGAAGAGTCTCTGGTTTGGAACTTTCATCAGCTGAGAGAGG 1591
Db 1754 AGTCTCTCCCGAGCTGGCAAAACGCTCTCTGGTTTGGGAACTTTCATCTCTTGGACAAAG 1813
Qy 1592 AGGAGCAGATCTTCTGGTTCATCAAGACAAACCTCTGAGCTCCATCAAGGCTGACATCG 1651
Db 1814 AAGAACAAATATTCTCTGCTGCTAAAGGACAAACCTCTCAGAGCATCAAGAGACATCG 1873
Qy 1652 TGACAGCTTCTCTGCTGCTTCCAGTCTCAGGACACAGCGTCTCTCCCAACAGAGCTTCC 1711
Db 1874 TCCATGCTTCTGCTGCTTCCAGGCTGAGTTCAGAGTGTGCTGTCACAGCCAGCTTCA 1933

1712	QY	GGGCCGAGTACAAAGGCCACGGGGGGGCGACCGCTGTTCCAGAACCGCGGTCAAGTTCCAGG	1771
1934	Db	GGGCCGAGTACAAAGGCCAGTGGCGGCCCTCCGCTTTCAAAGACCGCGTCCGCTTCCAGG	1993
1772	QY	TTGATATACCTACACGAGGGT-----GGGGAGGCGCAAGAGGAGA	1813
1994	Db	TGACATCAGTCCCTGTGAGGGTCCAGAGCCCTCCCGCGACGGGACGGCGGAGGTG	2053
1814	QY	ACGGCATCTACTCCGTCACTTCAACCCTGTCTCAGGCCCCAGCGCTCGCTTCAAGAGGG	1873
2054	Db	GTGGCATCTACTCCGTCACTTCACTCTCATCTCGGGTCCAGCCGTCGGTTCAAGCGAG	2113
1874	QY	TGCTGAGAGCCATCCAGGCCCGAGTGTCTGAGCACACAGCACCGCCCTGCGGCCCGCACT	1933
2114	Db	TGCTGAGAGCCATCCAGGCCAGCTCCTGAGCACTCATGACCGCCCTCCGTGCGAGGCC	2173
1934	QY	TGTCAGACACCACTAACTG	1952
2174	Db	TGGCAGACGAGAAGACGG	2192

RESULT 13

```

US-10-354-358-91
; Sequence 91, Application US/10354358
; Publication No. US20030157082A1
; GENERAL INFORMATION:
; APPLICANT: Millennium Pharmaceuticals, Inc
; APPLICANT: Hunter, John Joseph
; APPLICANT: MacBeth, Kyle J.
; APPLICANT: Tsai, Fong-ying
; APPLICANT: Lesoon, Andrea
; APPLICANT: Lightcap, Eric S.
; APPLICANT: Williamson, Mark
; APPLICANT: Rudolph-Owen, Laura A.
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR TREATING
; TITLE OF INVENTION: CANCER USING 140, 1470, 1686, 2089, 2427, 3702, 5891, 6428,
; TITLE OF INVENTION: 7181, 7660, 25641, 69583, 49863, 8897, 1682, 17667, 9235,
; TITLE OF INVENTION: 3703, 14171, 10359, 1660, 1450, 18894, 2088, 32427, 2160,
; TITLE OF INVENTION: 9252, 9389, 1642, 85269, 10297, 1584, 9525, 14124, 4469,
; TITLE OF INVENTION: 8990, 2100, 9288, 64698, 10480, 20893, 33230, 1586, 9943,
; TITLE OF INVENTION: 16334, 68862, 9011, 14031, 6178, 21225, 1420, 32236, 2099,
; TITLE OF INVENTION: 2150, 26583, 2784, 8941, 9811, 27444, 50566 OR 66428 MOLECULES
; FILE REFERENCE: MPI02-020PRNOMNIM
; CURRENT APPLICATION NUMBER: US/10/354,358
; CURRENT FILING DATE: 2003-01-30
; PRIOR APPLICATION NUMBER: US 60/353,600
; PRIOR FILING DATE: 2002-01-31
; PRIOR APPLICATION NUMBER: US 60/364,517
; PRIOR FILING DATE: 2002-03-15
; PRIOR APPLICATION NUMBER: US 60/371,075
; PRIOR FILING DATE: 2002-04-09
; PRIOR APPLICATION NUMBER: US 60/371,507
; PRIOR FILING DATE: 2002-04-10
; PRIOR APPLICATION NUMBER: US 60/372,984
; PRIOR FILING DATE: 2002-04-16
; PRIOR APPLICATION NUMBER: US 60/374,194
; PRIOR FILING DATE: 2002-04-19
; PRIOR APPLICATION NUMBER: US 60/382,995
; PRIOR FILING DATE: 2002-05-24
; PRIOR APPLICATION NUMBER: US 60/385,023
; PRIOR FILING DATE: 2002-05-31
; PRIOR APPLICATION NUMBER: US 60/388,853
; PRIOR FILING DATE: 2002-06-14
; PRIOR APPLICATION NUMBER: US 60/389,395
; PRIOR FILING DATE: 2002-06-17
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 122
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 91
; LENGTH: 2980
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:

```



```
QY 781 CTAGAGCATTTCAGAAACACATATGTGTATATAGGGGCAAGATGAGCCCGAACC----- 836
Db 601 CTGGAGCAAAATTCAGAAACATCTTTGGTACCTTAGGCGGGAACACAGAGCCAGACCCGTCG 660
QY 837 --AGAGCAGCCATTCCTCGAAGGTGCAGATCGCTCGCTGCCAGCCTCGAGGACATC 894
Db 661 CTGGAGCAGCCCTGGCGCGCGGTAGCCATGCGAGCCTTGCCATCCAAAGGAGAGCTG 720
QY 895 GACCCGACGCTGTGAGCAGCATGCACTCACTGGGCTGCTTCGAGAGCCGCAACAAAGCTG 954
Db 721 GACCCGACGCTGTGAGCAGCATGCACTCACTGGGCTGCTTCGAGAGCCGCGAGAGGCTG 780
QY 955 CTGAGGACCTGCTGCTCGAGAGAGAAACAGGAGAGATGATTACTTCTCTCTCTG 1014
Db 781 CATCGGAGCTGCGCAGTGTAGAGAGAGAAACCAAGAAAGATGATATATTCTGCTTTTG 840
QY 1015 GACCGAAAGAAAGTATCCCGAGCAGGAGATGAGACCTGCCCGCCGGAACGAGATA 1074
Db 841 GATCGAAGAGCGGTATCCCAAGCTGTGAGAACAGGACCTGCCCGCGAATGATGTT 900
QY 1075 GACCTTCCCGAAGCGTGTGACTTCCCGATGCTGAACCGGCAAGCGGCGGCA 1134
Db 901 GACCCCGCGGAGCGTGTGATTCTCCCATGCTGAGCGGTCAAGCGGAGCGGACCA 960
QY 1135 GNACGCAATCCATGAGGTGCTAGCGTGACGGA-----CGGCGGCTCCCGGTG 1185
Db 961 GAGCGGAAGTCCATGGAATGCTGAGCATCACCGATGCCGGGGGTGTGGCTCCCGCTGTA 1020
QY 1186 CTTGCGCGCGGCGCATTTAGATGATGCCAGACGCGCAGAGGTCTCGGTCCATCAGCGGT 1245
Db 1021 CCCACCGAGCGGCTTGGAGATGCGCCAGACAGCAGAGATCCCGTAGGTCTAGTGA 1080
QY 1246 GCCTCTCAGGCTTTTCCACAGCCCACTCAGACGCCCCCGGG----- 1288
Db 1081 GCCTCAGGCTCTGCTCTCAGCCCTCTAAGCAGCCCCAAGAGTGTCTTTTCCCTTT 1140
QY 1289 ----- 1288
Db 1141 TCACGAGAGCGGGGCTGGAGATGAGGCTCGAGCGGGGGCTCCCCGACTTCACAAACG 1200
QY 1289 -----TGACCCCTCACCCTC 1302
Db 1201 CAGACGCTGCTTCTCGGGGCCCCAGGGGTGGGGGCGCGGGGAGCAGCCCGCGCCCCC 1260
QY 1303 TCACCAAGGGGAGTCCCTCTCCCT-----ACC 1329
Db 1261 AGTGCCCGCTCCACACCCCTGCGCGCCCCCAGGCTCCCCGCGCTCCTCTGGCGGGACC 1320
QY 1330 CCCAAGGGGACACTGTTCACAGCGCAAGAGAGAGCCGGCTGGCAGCGCCCAACCCGACG 1389
Db 1321 CCTTTGCACCTGCGCTCTGCACACGCCCCCGGGGCCAGTCCACCGGACCCCCGGGGACAACA 1380
QY 1390 CCCCCGCTCAGCCCC-----AGCGTGGAGGGGTGCCCTGGAGGGCGCGGTCAACTTCC 1443
Db 1381 CCACCCCCCAGCCCCGGGGTGGCTCGGGGAGCGGCTGGAGAGTGTCTCAACTCC 1440
QY 1444 ATCAAGAACAGTTTCTGGGCTCACCCCGCTTCCACCGCGGAAACTCAAGTTCCGACG 1503
Db 1441 ATCCGCAACAGTCTCGGGCTCCCTCGCTTTTACCGCGGCAAGATGACAGTCCCTACC 1500
QY 1504 CCGGAGAGATGTCCAACTTGACACAGAGTGTCTCCCGAGAGCTGGCGAAGAGTCTCTGG 1563
Db 1501 GCTGAGGAGATGTTCAGTTGACGCGCAGAGTCTCTCCCGGAGCTGGCAAAACGCTCTCTGG 1560
QY 1564 TTGGAAGACTTCATCAGCCTGGAGAGAGAGGAGAGATCTTCGTGGTTCATCAAGACAAA 1623
Db 1561 TTGGAAGACTTCATCTCTGGAGCAAGAGAGAGAAATATTCTCTGCTTAAGAGCAAA 1620
QY 1624 CTTCTGAGCTCCATCAAGGCTGACATCGTGACGCTTCTCTGTGATTTCCAGTCTCAGC 1683
Db 1621 CTTCTCAGCAGCATCAAGAGCAGACATCGTCCATGCTTCTGTGATCCCCAGGCTGTAGT 1680
QY 1684 CACAGCTCATCTCCCAACAGAGCTTCCGGGCGGAGTACAGGGCCACGGGGGGGCCAGCC 1743
```

```
Db 1681 CACAGTGTGTGTACAGAGACCAGTTTCAGGGCCGAGTACAAAGGCCAGTGGCGGCCCTCC 1740
QY 1744 GTGTTCCAGAAAGCGGTCAAGTTTCCAGTTTATATACCTTACACGCGAGGT----- 1794
Db 1741 GTCTTCCAAAGGCCGCTCCGCTTCCAGTGTGACATCAGTCTCTGTAGGGTCCAGAGCCC 1800
QY 1795 -----GGGAGGCGCAGAGGAGAACCGCATCTACTCTCCGTCACTTCCACCTCTC 1845
Db 1801 TCCCGCGCAGCGGACGCGCAGCGAGGTGTGGCATCTACTCCGTCACTTCACTCTCATC 1860
QY 1846 TCAGGCCCCGAGCGCTGCTTCAAGAGGGTGTGGAGACCATCCAGGCCAGCTGCTGAGC 1905
Db 1861 TCGGGTCCCCAGCGCTCGGTTCAAGCGAGTGTGGAGACCATCCAGGCACAGCTCTCTGAGC 1920
QY 1906 ACACAGCAGCCCGCTGGCGGCCAGCACTGTCTCAGACACCACTAACTG 1952
Db 1921 ACTCATGACCAAGCCCTCCGTGCAGGCCCTGGCAGACGAGAGAACGG 1967
```

RESULT 15

```
US-10-116-326-3
; Sequence 3, Application US/10116326
; Publication No. US20030166889A1
; GENERAL INFORMATION:
; APPLICANT: Turner, C. Alexander Jr.
; APPLICANT: Mathur, Brian
; APPLICANT: Fiddle, Carl Johan
; TITLE OF INVENTION: No. US20030166889A1 Human Kinases and Polynucleotides Encoding t
; FILE REFERENCE: LEX-0332-USA
; CURRENT FILING DATE: 2002-04-04
; PRIOR FILING DATE: 2002-04-04
; PRIOR FILING DATE: 2001-04-06
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 2289
; TYPE: DNA
; ORGANISM: homo sapiens
US-10-116-326-3
```

Query Match 38.7%; Score 777.2; DB 12; Length 2289;

Best Local Similarity 67.9%; Pred. No. 2.5e-200;

Matches 1277; Conservative 0; Mismatches 408; Indels 195; Gaps 6;

```
QY 268 TATTTGTACTGTGTGTAGAACACGCTGCTAGGTGGTGTGAGCTTTCGACTACCTGGTGAAG 327
Db 265 TTTAGGTACTGTGTCTTGGAGACGCTCTCGGGGGTGTGAGCTATTCGACTACCTGGTAAAG 324
QY 328 AAGGGAGGCTGACGCTTAAGGAGGCTCGGAAGTTCTTCCGGCAGATCATCTCTGCGCTG 387
Db 325 AAGGGAGAGACTGACGCTCAAGAGGCGCCGAAAGTTCTTCCGCCAGATTGTGTCTGCGCTG 384
QY 388 GACTTCTGCCACAGCCTCATATGCCCAGAGGATCTGAACCTGAAACCTCTGCTG 447
Db 385 GACTTCTGCCACAGCTACTCATCTGCCCACAGACCTTAAAGCCCGAGAACCTGCTTTTG 444
QY 448 GACGAGAAGAACAACTCCCGCATCGCAGACTTTTGGCATGGCTCCCTGCGAGTTGGCGAC 507
Db 445 GATGAGAAAACAACTCCCGCATTCGAGACTTCGGCATGGCGTCCCTGCGAGTTGGGGAC 504
QY 508 AGCTGTGTGAGACCAAGCTGTGGGTCCCCCCTTACGCTGCCCGAGGATGATCCCGGGG 567
Db 505 AGCTCTCGAGACCAAGCTGCGGTCCCCCAATTATCGGTGTCCAGAGGTGATTAAGGGG 564
QY 568 GAGAGATATGACGCCCGAAGCGGAGCTGTGAGCTGCGCGCTCATCTGTTCGCTTG 627
Db 565 GAAAAATATGATGCGCCCGCGGAGACATGTGGAGCTGTGGAGTCTCTCTTTCGCTG 624
QY 628 CTGCTGGGGGCTGCGCTTCGAGATGACAACTTTCGACAGCTGCTGGAGAGGTGAAG 687
Db 625 CTGCTGGGGGCTGTGCTTTGATGACGACAACTCTCGCAGCTGCTGGAGAGGTGAAG 684
```

QY 688 CGGGCGTGTTCACATGCGCGACTTTATCCCGCCGACTGCGAGAGTCTGCTACGGGGC 747
DB 685 CGGGCGTGTTCACATGCGCGACTTTATCCCGCCGACTGCGAGAGTCTGCTACGGGGC 744
QY 748 ATGATCGAGGTGAGCGCGCGACCGCGCTCAGCGTAGAGCAGATTCAGAAACATATGG 807
DB 745 ATGATCGAGGTGAGCG 804
QY 808 TATATAGGGGCAAGATGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 861
DB 805 TACCTAGCGGGGCAAGATGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 864
QY 862 CAGATCCGCTGCTGCG 921
DB 865 GCCATCGGAGCGCTGCCATCCACCGGAGCGTGGACCGCGCGCGCGCGCGCGCGCGCGCG 924
QY 922 TCATCGGCTGCTTCCGAGCG 981
DB 925 TCATCGGCTGCTTCCGAGCG 984
QY 982 AACGAGAGAGATGATTTACTTCTCTCTGAGCGGCGCGCGCGCGCGCGCGCGCGCGCGCG 1041
DB 985 AACGAGAGAGATGATTTACTTCTCTCTGAGCGGCGCGCGCGCGCGCGCGCGCGCGCGCG 1044
QY 1042 GAGGATGAGGACCTTCCCG 1101
DB 1045 GAGGATGAGGACCTTCCCG 1104
QY 1102 CCGATGCTGAACCG 1161
DB 1105 CCGATGCTGAACCG 1164
QY 1162 GTGAC-----GGACCGCGCGCTCCCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1212
DB 1165 ATCACCAGTCCGGGGTGTGGTCCCTCTGACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1224
QY 1213 CAGCAGCGCGAGAGTCTCGGTGCAATCAGCGGTGCTCTCAGCGCGCGCGCGCGCGCGCGCGCG 1272
DB 1225 CAGCAGCGCGAGAGTCTCGGTGCAATCAGCGGTGCTCTCAGCGCGCGCGCGCGCGCGCGCGCG 1284
QY 1273 CTGACG 1286
DB 1285 CTAAGCAGCG 1344
QY 1287 ----- 1286
DB 1345 GCTCGAGCGGGGGTCCCG 1404
QY 1287 -----GGTGAACCGCTCAGCG 1326
DB 1405 GGTGGGGCGCGCGGGGAGCAGCG 1464
QY 1327 -----ACCG 1356
DB 1465 CCCCCAGGGTCCCG 1524
QY 1357 AAGGAGAGCGCGCGTGGCAGCG 1410
DB 1525 CGGGCGAGTCCACCGGGGACCGCGGGGACACACACCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1584
QY 1411 GAGGGGTGCGTGGAGGGCGCGGTCAATCTCATCAAGAACAGCTTCTGGGGTCAACC 1470
DB 1585 GGGGGAGCGCGCTGGAGGAGTCTCTCAACTCTCATCCGCAACAGCTTCTGGGGTCCCGCT 1644
QY 1471 CGCTTCCACCGCGGGAAGTCCGAGCGCGCGGAGGATGTCCTCAACCTGACACCA 1530
DB 1645 CGCTTCCACCGCGGGAAGTCCGAGCGCGGAGGATGTCCTCAACCTGACACCA 1704
QY 1531 GAGTCTGCCAGAGCTGGCGAAGAGTCTCTGGTTGGGAACTTCATCAGCGCTGGAGAG 1590
DB 1705 GAGTCTGCCAGAGCTGGCGAAGAGTCTCTGGTTGGGAACTTCATCTCTTGGACAAA 1764

QY 1591 GAGGAGCAGATCTTCGTGGTCAATCAAAAGACAAACCTCTGAGCTCCATCAAGGCTGACATC 1650
DB 1765 GAAGAAACAAATATCTCTGCTGCTAAAGGACAAACCTCTCAGCAGCATCAAGACGACATC 1824
QY 1651 GTGCAAGCGCTTCTGCTGATTCCTGATTCCTGATTCCTGATTCCTGATTCCTGATTCCTGAT 1710
DB 1825 GTCCATGCTTCTGCTGATTCCTGATTCCTGATTCCTGATTCCTGATTCCTGATTCCTGATTC 1884
QY 1711 CGGGCGCGAGTACAAGGCCACCGGGGGGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1770
DB 1885 AGGGCGCGAGTACAAGGCCACCGGGGGGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1944
QY 1771 GTTGATATCACTACACGGAGGT-----GGGGAGCGCGAGAGGAG 1812
DB 1945 GTGGACATCAGCTCTCTGAGGGTCCAGAGCGCTCCCGCGCGCGCGCGCGCGCGCGCGCG 2004
QY 1813 AACGGCATCTACTCCGTGATCTTCACTTCTCAGGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1872
DB 2005 GGTGGCATCTACTCCGTGATCTTCACTTCTCAGGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 2064
QY 1873 GTGGTGGAGACCATCCAGGCCCG 1932
DB 2065 GTGGTGGAGACCATCCAGGCCCG 2124
QY 1933 TTGTGACACCACTTACTG 1952
DB 2125 CTGGCAGACGAGAAGAACGG 2144

Search completed: November 27, 2003, 15:02:39
Job time : 1179.25 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 26, 2003, 12:01:57 ; Search time 28.2696 Seconds
(without alignments)
4358.326 Million cell updates/sec

Title: US-10-054-579-2

Perfect score: 3497

Sequence: 1 MTSTGKGGAHAQYGVPR.....TNCMMMTGRSLKCGIIPKS 668

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 673684 seqs, 184443283 residues

Total number of hits satisfying chosen parameters: 673684

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA:*
1: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep.*
2: /cgn2_6/ptodata/2/pubpaa/FCT_NEW_PUB.pep.*
3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep.*
4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep.*
5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep.*
6: /cgn2_6/ptodata/2/pubpaa/FCTUS_PUBCOMB.pep.*
7: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep.*
8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep.*
9: /cgn2_6/ptodata/2/pubpaa/US09A_PUBCOMB.pep.*
10: /cgn2_6/ptodata/2/pubpaa/US09B_PUBCOMB.pep.*
11: /cgn2_6/ptodata/2/pubpaa/US09C_PUBCOMB.pep.*
12: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pep.*
13: /cgn2_6/ptodata/2/pubpaa/US10A_PUBCOMB.pep.*
14: /cgn2_6/ptodata/2/pubpaa/US10B_PUBCOMB.pep.*
15: /cgn2_6/ptodata/2/pubpaa/US10C_PUBCOMB.pep.*
16: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pep.*
17: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep.*
18: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	3497	100.0	668	US-10-054-579-2	Sequence 2, Appli
2	3497	100.0	668	US-10-195-072-2	Sequence 2, Appli
3	3497	100.0	668	US-10-195-071-2	Sequence 2, Appli
4	3423	97.9	664	US-10-288-798-18	Sequence 18, Appli
5	3380	96.7	674	US-10-283-247-2	Sequence 2, Appli
6	3374	96.5	674	US-03-842-582-2	Sequence 2, Appli
7	3374	96.5	674	US-10-283-247-7	Sequence 7, Appli
8	3374	96.5	674	US-10-283-247-8	Sequence 8, Appli
9	3368	91.2	608	US-10-054-579-4	Sequence 4, Appli
10	3065	87.6	614	US-10-283-247-9	Sequence 9, Appli
11	3065	87.6	614	US-10-283-247-10	Sequence 10, Appli
12	3050	87.2	636	US-10-283-247-5	Sequence 5, Appli
13	3028	86.6	585	US-10-195-072-4	Sequence 4, Appli
14	3028	86.6	585	US-10-195-071-4	Sequence 4, Appli
15	3015	86.2	603	US-10-195-072-17	Sequence 17, Appli

16	3015	86.2	603	15	US-10-195-071-17	Sequence 17, Appli
17	2799	80.0	778	12	US-10-354-358-92	Sequence 92, Appli
18	2799	80.0	778	12	US-10-116-326-2	Sequence 2, Appli
19	2545.5	72.8	703	12	US-10-116-326-6	Sequence 6, Appli
20	2386.5	68.2	762	12	US-10-116-326-4	Sequence 4, Appli
21	796.5	22.8	745	12	US-10-161-565-24	Sequence 24, Appli
22	796.5	22.8	745	15	US-10-195-101-36	Sequence 36, Appli
23	794.5	22.7	691	10	US-09-919-585-6	Sequence 6, Appli
24	794.5	22.7	724	10	US-09-919-585-9	Sequence 9, Appli
25	790.5	22.6	722	8	US-08-817-832B-32	Sequence 32, Appli
26	790	22.6	1518	10	US-09-801-368-152	Sequence 152, App
27	781	22.3	744	10	US-09-835-081-4	Sequence 4, Appli
28	779.5	22.3	752	10	US-09-835-081-2	Sequence 2, Appli
29	778.5	22.3	688	12	US-10-161-565-28	Sequence 28, Appli
30	778.5	22.3	688	12	US-10-161-565-29	Sequence 29, Appli
31	777.5	22.2	795	10	US-09-919-585-12	Sequence 12, Appli
32	777.5	22.2	795	12	US-10-161-565-25	Sequence 25, Appli
33	777.5	22.2	795	15	US-10-142-356-9	Sequence 9, Appli
34	771.5	22.1	783	9	US-09-815-915-2	Sequence 2, Appli
35	771.5	22.1	783	12	US-10-393-316-2	Sequence 2, Appli
36	769	22.0	1369	12	US-10-288-798-7	Sequence 7, Appli
37	768.5	22.0	729	12	US-10-161-565-26	Sequence 26, Appli
38	768.5	22.0	729	15	US-10-142-356-11	Sequence 11, Appli
39	768.5	22.0	729	15	US-10-195-101-33	Sequence 33, Appli
40	765.5	21.9	783	11	US-09-823-187-90	Sequence 90, Appli
41	764.5	21.9	793	15	US-10-195-101-32	Sequence 32, Appli
42	762.5	21.8	744	10	US-09-919-585-3	Sequence 3, Appli
43	759.5	21.7	713	12	US-10-161-565-27	Sequence 27, Appli
44	754	21.6	786	11	US-09-823-187-91	Sequence 91, Appli
45	749	21.4	776	11	US-09-823-187-92	Sequence 92, Appli

ALIGNMENTS

RESULT 1

US-10-054-579-2
; Sequence 2, Application US/10054579
; Publication No. US20020137913A1
; GENERAL INFORMATION:
; APPLICANT: Turner, C. Alexander Jr.
; APPLICANT: Mathur, Brian
; TITLE OF INVENTION: No. US20020137913A1el Human Kinases and Polynucleotides Encoding t
; FILE REFERENCE: LEX-0300-USA
; CURRENT APPLICATION NUMBER: US/10/054.579
; CURRENT FILING DATE: 2002-01-22
; PRIOR APPLICATION NUMBER: US 60/263,378
; PRIOR FILING DATE: 2001-01-23
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 668
; TYPE: PRT
; ORGANISM: homo sapiens
US-10-054-579-2

Query Match	100.0%;	Score 3497;	DB 14;	Length 668;
Best Local Similarity	100.0%;	Pred. No. 2.8e-218;		
Matches	668;	Conservative	0;	Mismatches 0; Indels 0; Gaps 0;
Qy	1	MTSTGKGGAHAQYGVPRLEKTLGQGLVKGHVCTCQKVAIKIVNREKLSVVL	60	
Db	1	MTSTGKGGAHAQYGVPRLEKTLGQGLVKGHVCTCQKVAIKIVNREKLSVVL	60	
Qy	61	MKVERETAILKLIIEHPVHLKLDVYENKYLVLVLEHVSGGELFDYLVKKGLTPKEARK	120	
Db	61	MKVERETAILKLIIEHPVHLKLDVYENKYLVLVLEHVSGGELFDYLVKKGLTPKEARK	120	
Qy	121	FFRQIISALDFCHSHSICHRDLKPELNLDEKKNIRIADFGMASIQVGDLSLETSCGSPH	180	
Db	121	FFRQIISALDFCHSHSICHRDLKPELNLDEKKNIRIADFGMASIQVGDLSLETSCGSPH	180	
Qy	181	YACPEVIRGEKYDGRKADWSCGVILFALLVGLPFDDNLRQLLEKVKRGVFMHPHIP	240	

Db 181 YACPEVIRGEKYDGRKADVWSCGVILFALLVGLPFDNDNLRQLLEKVKRGVFMHPFIP 240
Qy 241 PDCOSLARGMEVDAARRLTLEHIQKHIWYIGGKNEPEPEQIPRKVQIIRSLPSLEIDP 300
Db 241 PDCOSLARGMEVDAARRLTLEHIQKHIWYIGGKNEPEPEQIPRKVQIIRSLPSLEIDP 300
Qy 301 DVLDSMHSGLGCFDRNRKLLQDLSSEENQEKMIYFLLLDKERYPSQDEDELPNRNEIDP 360
Db 301 DVLDSMHSGLGCFDRNRKLLQDLSSEENQEKMIYFLLLDKERYPSQDEDELPNRNEIDP 360
Qy 361 PRKRVDSFMLNRHGRKRRPERKSMVLSVTDGGSVPARRAIEMAQHQQRSRISGASSGL 420
Db 361 PRKRVDSFMLNRHGRKRRPERKSMVLSVTDGGSVPARRAIEMAQHQQRSRISGASSGL 420
Qy 421 STSPLSSPRVTPHPSPRGSPPLTPKGTVPHTPKESPAGTNPPTPSSPSVGGVPMWRLN 480
Db 421 STSPLSSPRVTPHPSPRGSPPLTPKGTVPHTPKESPAGTNPPTPSSPSVGGVPMWRLN 480
Qy 481 SIKNSFLGSPRFRHRRKLVQPTPEMSNLTPSSPELAKKSWFGNFIISLEKEEQIFVVIKD 540
Db 481 SIKNSFLGSPRFRHRRKLVQPTPEMSNLTPSSPELAKKSWFGNFIISLEKEEQIFVVIKD 540
Qy 541 KPLSSIKADIIVHAFSLPSLSHVSISQTSFRAEYKATGGPAVFOKPKVQVDITYTEGGE 600
Db 541 KPLSSIKADIIVHAFSLPSLSHVSISQTSFRAEYKATGGPAVFOKPKVQVDITYTEGGE 600
Qy 601 AQKENGIVSYVFTLLSGSPRRFRKRVVETIQALLSTHDPAAQHLSDTTNCMENMTGRLS 660
Db 601 AQKENGIVSYVFTLLSGSPRRFRKRVVETIQALLSTHDPAAQHLSDTTNCMENMTGRLS 660
Qy 661 KCGIIPKS 668
Db 661 KCGIIPKS 668

RESULT 2

US-10-195-072-2
; Sequence 2, Application US/10195072
; Publication No. US20030092036A1
; GENERAL INFORMATION:
; APPLICANT: Origene Technologies
; TITLE OF INVENTION: Full-Length Serine Protein Kinase in Brain and Pancreas
; FILE REFERENCE: 16U 101 C2
; CURRENT APPLICATION NUMBER: US/10/195,072
; CURRENT FILING DATE: 2002-07-15
; PRIOR APPLICATION NUMBER: US 09/930,181
; PRIOR FILING DATE: 2001-08-16
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: Patent in version 3.1
; SEQ ID NO 2
; LENGTH: 668
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-195-072-2

Query Match 100.0%; Score 3497; DB 15; Length 668;
Best Local Similarity 100.0%; Pred. No. 2.8e-218;
Matches 668; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MTSTGKGGGAHQAYVGPYRLEKTLGKGTGLVKGVCVTCQKVAIKIVNREKLSSEVL 60
Db 1 MTSTGKGGGAHQAYVGPYRLEKTLGKGTGLVKGVCVTCQKVAIKIVNREKLSSEVL 60
Qy 61 MKVEREAILKLIIEHPHVLKLDVYENKKYLYLVLEHVS GGELFDYLVKKGLRTPKEARK 120
Db 61 MKVEREAILKLIIEHPHVLKLDVYENKKYLYLVLEHVS GGELFDYLVKKGLRTPKEARK 120
Qy 121 FFRQIISALDFCHSHSICHRLDKPENLLLDKKNIRIADFGWASIQVGDLSLETSCGSPH 180
Db 121 FFRQIISALDFCHSHSICHRLDKPENLLLDKKNIRIADFGWASIQVGDLSLETSCGSPH 180
Qy 181 YACPEVIRGEKYDGRKADVWSCGVILFALLVGLPFDNDNLRQLLEKVKRGVFMHPFIP 240

Db 181 YACPEVIRGEKYDGRKADVWSCGVILFALLVGLPFDNDNLRQLLEKVKRGVFMHPFIP 240
Qy 241 PDCOSLARGMEVDAARRLTLEHIQKHIWYIGGKNEPEPEQIPRKVQIIRSLPSLEIDP 300
Db 241 PDCOSLARGMEVDAARRLTLEHIQKHIWYIGGKNEPEPEQIPRKVQIIRSLPSLEIDP 300
Qy 301 DVLDSMHSGLGCFDRNRKLLQDLSSEENQEKMIYFLLLDKERYPSQDEDELPNRNEIDP 360
Db 301 DVLDSMHSGLGCFDRNRKLLQDLSSEENQEKMIYFLLLDKERYPSQDEDELPNRNEIDP 360
Qy 361 PRKRVDSFMLNRHGRKRRPERKSMVLSVTDGGSVPARRAIEMAQHQQRSRISGASSGL 420
Db 361 PRKRVDSFMLNRHGRKRRPERKSMVLSVTDGGSVPARRAIEMAQHQQRSRISGASSGL 420
Qy 421 STSPLSSPRVTPHPSPRGSPPLTPKGTVPHTPKESPAGTNPPTPSSPSVGGVPMWRLN 480
Db 421 STSPLSSPRVTPHPSPRGSPPLTPKGTVPHTPKESPAGTNPPTPSSPSVGGVPMWRLN 480
Qy 481 SIKNSFLGSPRFRHRRKLVQPTPEMSNLTPSSPELAKKSWFGNFIISLEKEEQIFVVIKD 540
Db 481 SIKNSFLGSPRFRHRRKLVQPTPEMSNLTPSSPELAKKSWFGNFIISLEKEEQIFVVIKD 540
Qy 541 KPLSSIKADIIVHAFSLPSLSHVSISQTSFRAEYKATGGPAVFOKPKVQVDITYTEGGE 600
Db 541 KPLSSIKADIIVHAFSLPSLSHVSISQTSFRAEYKATGGPAVFOKPKVQVDITYTEGGE 600
Qy 601 AQKENGIVSYVFTLLSGSPRRFRKRVVETIQALLSTHDPAAQHLSDTTNCMENMTGRLS 660
Db 601 AQKENGIVSYVFTLLSGSPRRFRKRVVETIQALLSTHDPAAQHLSDTTNCMENMTGRLS 660
Qy 661 KCGIIPKS 668
Db 661 KCGIIPKS 668

RESULT 3

US-10-195-071-2
; Sequence 2, Application US/10195071
; Publication No. US20030096271A1
; GENERAL INFORMATION:
; APPLICANT: Origene Technologies
; TITLE OF INVENTION: Full-Length Serine Protein Kinase in Brain and Pancreas
; FILE REFERENCE: 16U 101 C1
; CURRENT APPLICATION NUMBER: US/10/195,071
; CURRENT FILING DATE: 2002-07-15
; PRIOR APPLICATION NUMBER: US 09/930,181
; PRIOR FILING DATE: 2001-08-16
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: Patent in version 3.1
; SEQ ID NO 2
; LENGTH: 668
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-195-071-2

Query Match 100.0%; Score 3497; DB 15; Length 668;
Best Local Similarity 100.0%; Pred. No. 2.8e-218;
Matches 668; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MTSTGKGGGAHQAYVGPYRLEKTLGKGTGLVKGVCVTCQKVAIKIVNREKLSSEVL 60
Db 1 MTSTGKGGGAHQAYVGPYRLEKTLGKGTGLVKGVCVTCQKVAIKIVNREKLSSEVL 60
Qy 61 MKVEREAILKLIIEHPHVLKLDVYENKKYLYLVLEHVS GGELFDYLVKKGLRTPKEARK 120
Db 61 MKVEREAILKLIIEHPHVLKLDVYENKKYLYLVLEHVS GGELFDYLVKKGLRTPKEARK 120
Qy 121 FFRQIISALDFCHSHSICHRLDKPENLLLDKKNIRIADFGWASIQVGDLSLETSCGSPH 180
Db 121 FFRQIISALDFCHSHSICHRLDKPENLLLDKKNIRIADFGWASIQVGDLSLETSCGSPH 180
Qy 181 YACPEVIRGEKYDGRKADVWSCGVILFALLVGLPFDNDNLRQLLEKVKRGVFMHPFIP 240

Db 181 YACFEVIREKDGKADVWSCGVILFALLVAGALPDDNNLQLEKVKRGVFMHPHIP 240
QY 241 PDCQSLRGMIEVDAARLTLEHIQKHIWYIGKNEPEPEQIPRKVQIRSLPSLEDIDP 300
Db 241 PDCQSLRGMIEVDAARLTLEHIQKHIWYIGKNEPEPEQIPRKVQIRSLPSLEDIDP 300
QY 301 DVLDSHSGCGRDRNKLQDILLSEENQEKMIYFLLDRKERYPSQDEDELPPEINIDP 360
Db 301 DVLDSHSGCGRDRNKLQDILLSEENQEKMIYFLLDRKERYPSQDEDELPPEINIDP 360
QY 361 PRKRVDSPLNLRHGRKRPRKSNVLSVTDGSPVPARRAIEMAHQGRSISGASSGL 420
Db 361 PRKRVDSPLNLRHGRKRPRKSNVLSVTDGSPVPARRAIEMAHQGRSISGASSGL 420
QY 421 STSPSSPRVTHPSRGSPLPTPKGTVPHTPKESPAGTNPPTPSSPSVGGVPMRRLN 480
Db 421 STSPSSPRVTHPSRGSPLPTPKGTVPHTPKESPAGTNPPTPSSPSVGGVPMRRLN 480
QY 481 SIKNSFLGSPRRHRRKLQVPTPEMSNLTPSSPELAKKSWFNFISLEKEBQIFVVIKD 540
Db 481 SIKNSFLGSPRRHRRKLQVPTPEMSNLTPSSPELAKKSWFNFISLEKEBQIFVVIKD 540
QY 541 KPLSSIKADIVAFSLPSLSHSVISQTSFRAEYKATGPAVFPKVPKQVDITTEGGE 600
Db 541 KPLSSIKADIVAFSLPSLSHSVISQTSFRAEYKATGPAVFPKVPKQVDITTEGGE 600
QY 601 AQKENGIVSVTFLSGPSRRPRKRVVETIQALLSTHDPAAQHLSDTTNCNEMMTGRLS 660
Db 601 AQKENGIVSVTFLSGPSRRPRKRVVETIQALLSTHDPAAQHLSDTTNCNEMMTGRLS 660
QY 661 KCGIIPKS 668
Db 661 KCGIIPKS 668

RESULT 4
US-10-288-798-18
; Sequence 18, Application US/10288798
; Publication No. US20030207299A1
; GENERAL INFORMATION:
; APPLICANT: BANDMAN, Olga; NGUYEN, Danniell B.;
; APPLICANT: WALIA, Narinder K.; HAFALIA, April J.A.;
; APPLICANT: YAO, Monique G.; GANDHI, Ameena R.;
; APPLICANT: GURURAJAN, Rajagopal; DING, Li;
; APPLICANT: PATTERSON, Chandra; YUE, Henry;
; APPLICANT: BAUGHN, Mariah R.; TRIBOULEY, Catherine M.;
; APPLICANT: THORNTON, Michael; ELLIOTT, Vicki S.;
; APPLICANT: LU, Yan; ISON, Craig H.;
; APPLICANT: AU-YOUNG, Janice; TANG, Y. Tom;
; APPLICANT: AZIMZAI, Yalda; BURRILL, John D.;
; APPLICANT: MARCUS, Gregory A.; ZINGLER, Kurt A.;
; APPLICANT: LU, Dying Aina M.; LAL, Preeti G.;
; APPLICANT: RAMKUMAR, Jayalaxmi; WARREN, Bridget A.;
; APPLICANT: KEARNEY, Liam; POLICKY, Jennifer L.;
; APPLICANT: THANGAVELU, Kavitha; BURFORD, Neil
; TITLE OF INVENTION: HUMAN KINASES
; FILE REFERENCE: PI-0209 USA
; CURRENT APPLICATION NUMBER: US/10/288,798
; CURRENT FILING DATE: 2002-11-01
; PRIOR APPLICATION NUMBER: PCT/US01/27219
; PRIOR FILING DATE: 2001-08-31
; PRIOR APPLICATION NUMBER: US 60/240,542
; PRIOR FILING DATE: 2000-10-13
; PRIOR APPLICATION NUMBER: US 60/238,389
; PRIOR FILING DATE: 2000-10-06
; PRIOR APPLICATION NUMBER: US 60/236,499
; PRIOR FILING DATE: 2000-09-29
; PRIOR APPLICATION NUMBER: US 60/234,902
; PRIOR FILING DATE: 2000-09-22
; PRIOR APPLICATION NUMBER: US 60/232,654
; PRIOR FILING DATE: 2000-09-14
; PRIOR APPLICATION NUMBER: US 60/231,357

; PRIOR FILING DATE: 2000-09-08
; PRIOR APPLICATION NUMBER: US 60/229,873
; FILING DATE: 2000-08-31
; NUMBER OF SEQ ID NOS: 48
; SOFTWARE: PERL Program
; SEQ ID NO 18
; LENGTH: 664
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: Incyte ID No. US20030207299A1 4022651CDI
US-10-288-798-18

Query Match 97.9%; Score 3423; DB 12; Length 664;
Best Local Similarity 100.0%; Pred. No. 1.7e-213;
Matches 654; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 15 YVGPVRLKTLGKGTGLVKGVLGHVCTVCQKVAIKIVNREKLSVLMKVEREIALKLE 74
Db 11 YVGPVRLKTLGKGTGLVKGVLGHVCTVCQKVAIKIVNREKLSVLMKVEREIALKLE 70
QY 75 HPHVLKLDHVVYENKYLVLVLEHVSQGLPDLVVKKGLRTPKEARKFPFQRIISALDFCHS 134
Db 71 HPHVLKLDHVVYENKYLVLVLEHVSQGLPDLVVKKGLRTPKEARKFPFQRIISALDFCHS 130
QY 135 HSIHRDLKPNLLDEKNIRIADFGWASIQVGSLLTSCGSPHYACPEVIRGEKYDG 194
Db 131 HSIHRDLKPNLLDEKNIRIADFGWASIQVGSLLTSCGSPHYACPEVIRGEKYDG 190
QY 195 RKADVWSCGVILFALLVAGALPDDNNLQLEKVKRGVFMHPHIPPPCOSLLRGMIEVD 254
Db 191 RKADVWSCGVILFALLVAGALPDDNNLQLEKVKRGVFMHPHIPPPCOSLLRGMIEVD 250
QY 255 AARLTLEHIQKHIWYIGKNEPEPEQIPRKVQIRSLPSLEDIDPDLDSHSGCGRFD 314
Db 251 AARLTLEHIQKHIWYIGKNEPEPEQIPRKVQIRSLPSLEDIDPDLDSHSGCGRFD 310
QY 315 RNKLQDILLSEENQEKMIYFLLDRKERYPSQDEDELPPEINIDPPEKRVDSPLNLRHG 374
Db 311 RNKLQDILLSEENQEKMIYFLLDRKERYPSQDEDELPPEINIDPPEKRVDSPLNLRHG 370
QY 375 KRRPERKSMEVLSVTDGSPVPARRAIEMAHQGRSISGASSGLSTPLSSPRVTHP 434
Db 371 KRRPERKSMEVLSVTDGSPVPARRAIEMAHQGRSISGASSGLSTPLSSPRVTHP 430
QY 435 SPRGSPLPTPKGTVPHTPKESPAGTNPPTPSSPSVGGVPMRRLNLSKNSFLGSPRFR 494
Db 431 SPRGSPLPTPKGTVPHTPKESPAGTNPPTPSSPSVGGVPMRRLNLSKNSFLGSPRFR 490
QY 495 RKLQVPTPEMSNLTPSSPELAKKSWFNFISLEKEBQIFVVIKDKPLSSIKADIVHAF 554
Db 491 RKLQVPTPEMSNLTPSSPELAKKSWFNFISLEKEBQIFVVIKDKPLSSIKADIVHAF 550
QY 555 LSIPSLSHSVISQTSFRAEYKATGPAVFPKVPKQVDITTEGGAOKENGIVSVTFTL 614
Db 551 LSIPSLSHSVISQTSFRAEYKATGPAVFPKVPKQVDITTEGGAOKENGIVSVTFTL 610
QY 615 LSGPSRRPRKRVVETIQALLSTHDPAAQHLSDTTNCNEMMTGRLSKCGIIPKS 668
Db 611 LSGPSRRPRKRVVETIQALLSTHDPAAQHLSDTTNCNEMMTGRLSKCGIIPKS 664

RESULT 5
US-10-283-247-2
; Sequence 2, Application US/10283247
; Publication No. US20030119037A1
; GENERAL INFORMATION:
; APPLICANT: NEELAM, Beena et al.
; TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC
; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES
; TITLE OF INVENTION: THEREOF
; FILE REFERENCE: CL001304

; CURRENT APPLICATION NUMBER: US/10/283,247
; CURRENT FILING DATE: 2002-10-30
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 674
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-283-247-2

Query Match 96.7%; Score 3380; DB 15; Length 674;
Best Local Similarity 99.8%; Pred. No. 1.1e-210;
Matches 646; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
Qy 1 MTSTKDGGAQAQYGVGYRLEKTLGKGTGLVGLGVHCVTCQKVAIKIVNREKLSVSL 60
Db 1 MTSTKDGGAQAQYGVGYRLEKTLGKGTGLVGLGVHCVTCQKVAIKIVNREKLSVSL 60
Qy 61 MKVERETAILKLEHPPHVLKLDVYENKKYLYLVLEHVSGLGELFDYLVKKGRLLTPKEARK 120
Db 61 MKVERETAILKLEHPPHVLKLDVYENKKYLYLVLEHVSGLGELFDYLVKKGRLLTPKEARK 120
Qy 121 FFRQIISALDFCHSHSICHRLDKPENLLDDEKNIRIADFGMASLQVGDLSLETSCGSPH 180
Db 121 FFRQIISALDFCHSHSICHRLDKPENLLDDEKNIRIADFGMASLQVGDLSLETSCGSPH 180
Qy 181 YACPEVIRGEKYDGRKADVWSCGVILFALLVGLPFDNDNLRQLLEKVKRGVFMHPHFIP 240
Db 181 YACPEVIRGEKYDGRKADVWSCGVILFALLVGLPFDNDNLRQLLEKVKRGVFMHPHFIP 240
Qy 241 PDCQSLLRGMIEVDAARLTLEHIQKHIWYIGGKNEPEPEQPIPRKQVIRSLPSLEDIDP 300
Db 241 PDCQSLLRGMIEVDAARLTLEHIQKHIWYIGGKNEPEPEQPIPRKQVIRSLPSLEDIDP 300
Qy 301 DVLDSMHSILGCFDRNKLQDLLSEBENQEKMIYFLLDRKERYPSQEDDLPPRNEIDP 360
Db 301 DVLDSMHSILGCFDRNKLQDLLSEBENQEKMIYFLLDRKERYPSQEDDLPPRNEIDP 360
Qy 361 PRKRVDSPLNLRHGKRRPERKSMVLSVTDGSPVPARRAIEMAHQGQRSISGASSGL 420
Db 361 PRKRVDSPLNLRHGKRRPERKSMVLSVTDGSPVPARRAIEMAHQGQRSISGASSGL 420
Qy 421 STSPLSSPRVTPHSPRGSPLTPKGTVPHTPKESPAGTNPPTSPSPSVGGVPMRRLN 480
Db 421 STSPLSSPRVTPHSPRGSPLTPKGTVPHTPKESPAGTNPPTSPSPSVGGVPMRRLN 480
Qy 481 SIKNSFLGSPFRHRRKLOVPTPEMSNLTPSSPELAKKSWFGNFI SLEKEEQIFVVIKD 540
Db 481 SIKNSFLGSPFRHRRKLOVPTPEMSNLTPSSPELAKKSWFGNFI SLEKEEQIFVVIKD 540
Qy 541 KPLSSIKADIVHAFSLSPSLSHSVISQTSFRAEYKATGPAVFOKPVKFOVDITYTTEGGE 600
Db 541 KPLSSIKADIVHAFSLSPSLSHSVISQTSFRAEYKATGPAVFOKPVKFOVDITYTTEGGE 600
Qy 601 AQKENGIVSYVFTLLSGSPRRFRKRVVETIQALLSTHDPAPAAQLSD 647
Db 601 AQKENGIVSYVFTLLSGSPRRFRKRVVETIQALLSTHDPAPAAQLSD 647

RESULT 6
US-09-842-582-2
; Sequence 2, Application US/09842582
; Patent No. US20020155570A1
; GENERAL INFORMATION:
; APPLICANT: Millennium Pharmaceuticals, Inc.
; APPLICANT: Meyers, Rachel
; TITLE OF INVENTION: 2246, NOVEL PROTEIN KINASE MOLECULES AND
; TITLE OF INVENTION: USES THEREFOR
; FILE REFERENCE: 38155-20054.00
; CURRENT APPLICATION NUMBER: US/09/842,582
; CURRENT FILING DATE: 2001-04-25
; PRIOR APPLICATION NUMBER: US 60/199,391
; PRIOR FILING DATE: 2000-04-25

; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 674
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-842-582-2

Query Match 96.5%; Score 3374; DB 10; Length 674;
Best Local Similarity 99.7%; Pred. No. 2.6e-210;
Matches 645; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
Qy 1 MTSTKDGGAQAQYGVGYRLEKTLGKGTGLVGLGVHCVTCQKVAIKIVNREKLSVSL 60
Db 1 MTSTKDGGAQAQYGVGYRLEKTLGKGTGLVGLGVHCVTCQKVAIKIVNREKLSVSL 60
Qy 61 MKVERETAILKLEHPPHVLKLDVYENKKYLYLVLEHVSGLGELFDYLVKKGRLLTPKEARK 120
Db 61 MKVERETAILKLEHPPHVLKLDVYENKKYLYLVLEHVSGLGELFDYLVKKGRLLTPKEARK 120
Qy 121 FFRQIISALDFCHSHSICHRLDKPENLLDDEKNIRIADFGMASLQVGDLSLETSCGSPH 180
Db 121 FFRQIISALDFCHSHSICHRLDKPENLLDDEKNIRIADFGMASLQVGDLSLETSCGSPH 180
Qy 181 YACPEVIRGEKYDGRKADVWSCGVILFALLVGLPFDNDNLRQLLEKVKRGVFMHPHFIP 240
Db 181 YACPEVIRGEKYDGRKADVWSCGVILFALLVGLPFDNDNLRQLLEKVKRGVFMHPHFIP 240
Qy 241 PDCQSLLRGMIEVDAARLTLEHIQKHIWYIGGKNEPEPEQPIPRKQVIRSLPSLEDIDP 300
Db 241 PDCQSLLRGMIEVDAARLTLEHIQKHIWYIGGKNEPEPEQPIPRKQVIRSLPSLEDIDP 300
Qy 301 DVLDSMHSILGCFDRNKLQDLLSEBENQEKMIYFLLDRKERYPSQEDDLPPRNEIDP 360
Db 301 DVLDSMHSILGCFDRNKLQDLLSEBENQEKMIYFLLDRKERYPSQEDDLPPRNEIDP 360
Qy 361 PRKRVDSPLNLRHGKRRPERKSMVLSVTDGSPVPARRAIEMAHQGQRSISGASSGL 420
Db 361 PRKRVDSPLNLRHGKRRPERKSMVLSVTDGSPVPARRAIEMAHQGQRSISGASSGL 420
Qy 421 STSPLSSPRVTPHSPRGSPLTPKGTVPHTPKESPAGTNPPTSPSPSVGGVPMRRLN 480
Db 421 STSPLSSPRVTPHSPRGSPLTPKGTVPHTPKESPAGTNPPTSPSPSVGGVPMRRLN 480
Qy 481 SIKNSFLGSPFRHRRKLOVPTPEMSNLTPSSPELAKKSWFGNFI SLEKEEQIFVVIKD 540
Db 481 SIKNSFLGSPFRHRRKLOVPTPEMSNLTPSSPELAKKSWFGNFI SLEKEEQIFVVIKD 540
Qy 541 KPLSSIKADIVHAFSLSPSLSHSVISQTSFRAEYKATGPAVFOKPVKFOVDITYTTEGGE 600
Db 541 KPLSSIKADIVHAFSLSPSLSHSVISQTSFRAEYKATGPAVFOKPVKFOVDITYTTEGGE 600
Qy 601 AQKENGIVSYVFTLLSGSPRRFRKRVVETIQALLSTHDPAPAAQLSD 647
Db 601 AQKENGIVSYVFTLLSGSPRRFRKRVVETIQALLSTHDPAPAAQLSD 647

RESULT 7
US-10-283-247-7
; Sequence 7, Application US/10283247
; Publication No. US20030119037A1
; GENERAL INFORMATION:
; APPLICANT: NEELAM, Beena et al.
; TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC
; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES
; TITLE OF INVENTION: THEREOF
; FILE REFERENCE: CL001304
; CURRENT APPLICATION NUMBER: US/10/283,247
; CURRENT FILING DATE: 2002-10-30
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 7
; LENGTH: 674

TYPE: PRT
ORGANISM: Homo sapiens
US-10-283-247-7

Query Match 96.5%; Score 3374; DB 15; Length 674;
Best Local Similarity 99.7%; Pred. No. 2.6e-210;
Matches 645; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 MTSTGKGGGAQAQYVGPYRLEKTLGKGTGLVGLGVHCVTCQKVAIKIVNREKLSSEVL 60
Db 1 MTSTGKGGGAQAQYVGPYRLEKTLGKGTGLVGLGVHCVTCQKVAIKIVNREKLSSEVL 60
Qy 61 MKVEREAILKLEIHPHVLKLDVYENKYLVLVLEHVSQGLDFDYLKKGRLTPKEARK 120
Db 61 MKVEREAILKLEIHPHVLKLDVYENKYLVLVLEHVSQGLDFDYLKKGRLTPKEARK 120
Qy 121 FFRQIISALDFCHSHSICHRDLKPNLLDEKNNIRIADFGWASLQVDSLLETSCGSPH 180
Db 121 FFRQIISALDFCHSHSICHRDLKPNLLDEKNNIRIADFGWASLQVDSLLETSCGSPH 180
Qy 181 YACPEVIRGEKYDGRKADWSCGVILFALLVGLPFDNNLRQLLEKVKRGVFMHPHPI 240
Db 181 YACPEVIRGEKYDGRKADWSCGVILFALLVGLPFDNNLRQLLEKVKRGVFMHPHPI 240
Qy 241 PDCQSLLRGMIEVDAAARLTLEHIQKHITWYIGGKNEPEPEQIPRKVQIRSLPSLEDIDP 300
Db 241 PDCQSLLRGMIEVDAAARLTLEHIQKHITWYIGGKNEPEPEQIPRKVQIRSLPSLEDIDP 300
Qy 301 DVLDSMHSILGCFRDRNKLQDLLSEENQEKMIYFLLLDRKERYPSQDEDLPPRNEIDP 360
Db 301 DVLDSMHSILGCFRDRNKLQDLLSEENQEKMIYFLLLDRKERYPSQDEDLPPRNEIDP 360
Qy 361 PRKRVDSPLNRHGKRRPERKSMEVLSVTDGSPVPARRAIEAQAQGRSISGASSGL 420
Db 361 PRKRVDSPLNRHGKRRPERKSMEVLSVTDGSPVPARRAIEAQAQGRSISGASSGL 420
Qy 421 STSPLSSPRVTPHSPRSGPLTPKGTVPVHTPKESPAGTNPPTPPSSPSVGVPMWRLN 480
Db 421 STSPLSSPRVTPHSPRSGPLTPKGTVPVHTPKESPAGTNPPTPPSSPSVGVPMWRLN 480
Qy 481 SIKNSFLGSPFRHRRKLVQVTPPEMSNLTPSSPELAKSWFGNFISLEKEQIFVVIKD 540
Db 481 SIKNSFLGSPFRHRRKLVQVTPPEMSNLTPSSPELAKSWFGNFISLEKEQIFVVIKD 540
Qy 541 KPLSSIKADIVHAFSLPSLSHVSISQTSFRAEYKATGPAVFOKPVKFQVDITYTGE 600
Db 541 KPLSSIKADIVHAFSLPSLSHVSISQTSFRAEYKATGPAVFOKPVKFQVDITYTGE 600
Qy 601 AQKENGIIYSVTFTLLSGPSRRFKRVVETIQALLSTHDPAAQHLSD 647
Db 601 AQKENGIIYSVTFTLLSGPSRRFKRVVETIQALLSTHDPAAQHLSE 647

RESULT 8

US-10-283-247-8
Sequence 8, Application US/10283247
Publication No. US20030119037A1
GENERAL INFORMATION:
APPLICANT: NEELAM, Beena et al.
TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC
TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES
TITLE OF INVENTION: THEREOF
FILE REFERENCE: CL001304
CURRENT APPLICATION NUMBER: US/10/283,247
CURRENT FILING DATE: 2002-10-30
NUMBER OF SEQ ID NOS: 10
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 8
LENGTH: 674
TYPE: PRT
ORGANISM: Homo sapiens
US-10-283-247-8

Query Match 96.5%; Score 3374; DB 15; Length 674;
Best Local Similarity 99.7%; Pred. No. 2.6e-210;
Matches 645; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 MTSTGKGGGAQAQYVGPYRLEKTLGKGTGLVGLGVHCVTCQKVAIKIVNREKLSSEVL 60
Db 1 MTSTGKGGGAQAQYVGPYRLEKTLGKGTGLVGLGVHCVTCQKVAIKIVNREKLSSEVL 60
Qy 61 MKVEREAILKLEIHPHVLKLDVYENKYLVLVLEHVSQGLDFDYLKKGRLTPKEARK 120
Db 61 MKVEREAILKLEIHPHVLKLDVYENKYLVLVLEHVSQGLDFDYLKKGRLTPKEARK 120
Qy 121 FFRQIISALDFCHSHSICHRDLKPNLLDEKNNIRIADFGWASLQVDSLLETSCGSPH 180
Db 121 FFRQIISALDFCHSHSICHRDLKPNLLDEKNNIRIADFGWASLQVDSLLETSCGSPH 180
Qy 181 YACPEVIRGEKYDGRKADWSCGVILFALLVGLPFDNNLRQLLEKVKRGVFMHPHPI 240
Db 181 YACPEVIRGEKYDGRKADWSCGVILFALLVGLPFDNNLRQLLEKVKRGVFMHPHPI 240
Qy 241 PDCQSLLRGMIEVDAAARLTLEHIQKHITWYIGGKNEPEPEQIPRKVQIRSLPSLEDIDP 300
Db 241 PDCQSLLRGMIEVDAAARLTLEHIQKHITWYIGGKNEPEPEQIPRKVQIRSLPSLEDIDP 300
Qy 301 DVLDSMHSILGCFRDRNKLQDLLSEENQEKMIYFLLLDRKERYPSQDEDLPPRNEIDP 360
Db 301 DVLDSMHSILGCFRDRNKLQDLLSEENQEKMIYFLLLDRKERYPSQDEDLPPRNEIDP 360
Qy 361 PRKRVDSPLNRHGKRRPERKSMEVLSVTDGSPVPARRAIEAQAQGRSISGASSGL 420
Db 361 PRKRVDSPLNRHGKRRPERKSMEVLSVTDGSPVPARRAIEAQAQGRSISGASSGL 420
Qy 421 STSPLSSPRVTPHSPRSGPLTPKGTVPVHTPKESPAGTNPPTPPSSPSVGVPMWRLN 480
Db 421 STSPLSSPRVTPHSPRSGPLTPKGTVPVHTPKESPAGTNPPTPPSSPSVGVPMWRLN 480
Qy 481 SIKNSFLGSPFRHRRKLVQVTPPEMSNLTPSSPELAKSWFGNFISLEKEQIFVVIKD 540
Db 481 SIKNSFLGSPFRHRRKLVQVTPPEMSNLTPSSPELAKSWFGNFISLEKEQIFVVIKD 540
Qy 541 KPLSSIKADIVHAFSLPSLSHVSISQTSFRAEYKATGPAVFOKPVKFQVDITYTGE 600
Db 541 KPLSSIKADIVHAFSLPSLSHVSISQTSFRAEYKATGPAVFOKPVKFQVDITYTGE 600
Qy 601 AQKENGIIYSVTFTLLSGPSRRFKRVVETIQALLSTHDPAAQHLSD 647
Db 601 AQKENGIIYSVTFTLLSGPSRRFKRVVETIQALLSTHDPAAQHLSE 647

RESULT 9

US-10-054-579-4
Sequence 4, Application US/10054579
Publication No. US20020137913A1
GENERAL INFORMATION:
APPLICANT: Turner, C. Alexander Jr.
APPLICANT: Mathur, Brian
TITLE OF INVENTION: No. US20020137913A1e1 Human Kinases and Polynucleotides Encoding t
FILE REFERENCE: LEX-0300-USA
CURRENT APPLICATION NUMBER: US/10/054,579
CURRENT FILING DATE: 2002-01-22
PRIOR APPLICATION NUMBER: US 60/263,378
PRIOR FILING DATE: 2001-01-23
NUMBER OF SEQ ID NOS: 4
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 4
LENGTH: 608
TYPE: PRT
ORGANISM: homo sapiens
US-10-054-579-4

Query Match 91.2%; Score 3188; DB 14; Length 608;
Best Local Similarity 100.0%; Pred. No. 2.5e-198;
Matches 608; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 241 PDCOSLLRGMIEVDAAARLTLLEHTQKHIWYIGKNEPEPEQPIPRKVQIRSLPSLEIDP 300
DB 181 PDCOSLLRGMSEVDAARLTLLEHTQKHIWYIGKNEPEPEQPIPRKVQIRSLPSLEIDP 240
QY 301 DVLDSMHSGLGCFDRNKLQDLSSEENQEKMIYFLLLDRKERYPSQDEDLPRNEIDP 360
DB 241 DVLDSMHSGLGCFDRNKLQDLSSEENQEKMIYFLLLDRKERYPSQDEDLPRNEIDP 300
QY 361 PRKRVDSPLNHRHGRKRRPERKSMELSVTDGSGVPARRAIEMAHQQRSSISGASSGL 420
DB 301 PRKRVDSPLNHRHGRKRRPERKSMELSVTDGSGVPARRAIEMAHQQRSSISGASSGL 360
QY 421 STSLSPSRVTPHPSRGSPPLTPKGTVPVHTPKESPAGTNPPTPPSPSGVGPWRARLN 480
DB 361 STSLSPSRVTPHPSRGSPPLTPKGTVPVHTPKESPAGTNPPTPPSPSGVGPWRARLN 420
QY 481 SIKNSFLGSPFRHRRKLQVPTPEMSNLTPSSPELAKSWFGNFI SLEKEEQIFVVIKD 540
DB 421 SIKNSFLGSPFRHRRKLQVPTPEMSNLTPSSPELAKSWFGNFI SLEKEEQIFVVIKD 480
QY 541 KPLSSIKADI VHAFLSIPSLSHSVISQTSFRAEYKATGGPAVFQKPVKFQVDITYTEGGE 600
DB 481 KPLSSIKADI VHAFLSIPSLSHSVISQTSFRAEYKATGGPAVFQKPVKFQVDITYTEGGE 540
QY 601 AOKENGISYVTFLLSGPSRRFRKRVETIQALLSTHDPAAQHLSD 647
DB 541 AOKENGISYVTFLLSGPSRRFRKRVETIQALLSTHDPAAQHLSE 587

RESULT 12
US-10-283-247-5
; Sequence 5, Application US/10283247
; Publication No. US20030113037A1
; GENERAL INFORMATION:
; APPLICANT: NEELAM, Bena et al.
; TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC
; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES
; TITLE OF INVENTION: THEREOF
; FILE REFERENCE: CL001304
; CURRENT APPLICATION NUMBER: US/10/283,247
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 5
; LENGTH: 636
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-283-247-5

Query Match 87.2%; Score 3050; DB 15; Length 636;
Best Local Similarity 96.2%; Pred. No. 2.3e-189;
Matches 586; Conservative 1; Mismatches 0; Indels 22; Gaps 1;
QY 61 MKVERETAILKLIIEHPVHLKLDVYENKXYLYLVLEHVSGLGELFDYLVKKGRLTPKARK 120
DB 1 MKVERETAILKLIIEHPVHLKLDVYENKXYLYLVLEHVSGLGELFDYLVKKGRLTPKARK 60
QY 121 FFRQIIISALDFCHSHSICHRLDKPENLLDDEKNIRIADFGMASLQVDSLLETSCGSPH 180
DB 61 FFRQIIISALDFCHSHSICHRLDKPENLLDDEKNIRIADFGMASLQVDSLLETSCGSPH 120
QY 181 YACPEVIRGEKYDGRKADVWSCGVILFALLVGLPFDNLRQLLEKVKRGVFMHPHFIP 240
DB 121 YACPEVIRGEKYDGRKADVWSCGVILFALLVGLPFDNLRQLLEKVKRGVFMHPHFIP 180
QY 241 PDCOSLLRGMIEVDAAARLTLLEHTQKHIWYIGKNEPEPEQPIPRKVQIRSLPSLEIDP 300
DB 181 PDCOSLLRGMIEVDAAARLTLLEHTQKHIWYIGKNEPEPEQPIPRKVQIRSLPSLEIDP 240
QY 301 DVLDSMHSGLGCFDRNKLQDLSSEENQEKMIYFLLLDRKERYPSQDEDLPRNEIDP 360
DB 241 DVLDSMHSGLGCFDRNKLQDLSSEENQEKMIYFLLLDRKERYPSQDEDLPRNEIDP 300

QY 361 PRKRVDSPLNHRHGRKRRPERKSMELSVTDGSGVPARRAIEMAHQHQ----- 408
DB 301 PRKRVDSPLNHRHGRKRRPERKSMELSVTDGSGVPARRAIEMAHQHQSKAMFSSKSLDIA 360
QY 409 -----RSRSISGASSGLSTSLSPSRVTPHPSRGSPPLTPKGTVPVHTPKESPAG 458
DB 361 EAHQPSKEDRSRSISGASSGLSTSLSPSRVTPHPSRGSPPLTPKGTVPVHTPKESPAG 420
QY 459 TPNPTPPSPSGVGPWRARLNSIKNSFLGSPFRHRRKLQVPTPEMSNLTPSSPELAK 518
DB 421 TPNPTPPSPSGVGPWRARLNSIKNSFLGSPFRHRRKLQVPTPEMSNLTPSSPELAK 480
QY 519 KSWFGNFI SLEKEEQIFVVIKOKPLSSIKADI VHAFLSIPSLSHSVISQTSFRAEYKATG 578
DB 481 KSWFGNFI SLEKEEQIFVVIKOKPLSSIKADI VHAFLSIPSLSHSVISQTSFRAEYKATG 540
QY 579 GPAVFQKPVKFQVDITYTEGGEAOKENGISYVTFLLSGPSRRFRKRVETIQALLSTHD 638
DB 541 GPAVFQKPVKFQVDITYTEGGEAOKENGISYVTFLLSGPSRRFRKRVETIQALLSTHD 600
QY 639 PPAHQHLSD 647
DB 601 PPAHQHLS 609

RESULT 13
US-10-195-072-4
; Sequence 4, Application US/10195072
; Publication No. US20030092036A1
; GENERAL INFORMATION:
; APPLICANT: Origene Technologies
; TITLE OF INVENTION: Full-Length Serine Protein Kinase in Brain and Pancreas
; FILE REFERENCE: 16U 101 C2
; CURRENT APPLICATION NUMBER: US/10/195,072
; CURRENT FILING DATE: 2002-07-15
; PRIOR APPLICATION NUMBER: US 09/930,181
; PRIOR FILING DATE: 2001-08-16
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 4
; LENGTH: 585
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-195-072-4

Query Match 86.6%; Score 3028; DB 15; Length 585;
Best Local Similarity 100.0%; Pred. No. 5.5e-188;
Matches 577; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 92 YLVLEHVSGLGELFDYLVKKGRLTPKARKFFRQIIISALDFCHSHSICHRLDKPENLLDDE 151
DB 9 YLVLEHVSGLGELFDYLVKKGRLTPKARKFFRQIIISALDFCHSHSICHRLDKPENLLDDE 68
QY 152 KNNIRIADFGMASLQVDSLLETSCGSPHYACPEVIRGEKYDGRKADVWSCGVILFALLV 211
DB 69 KNNIRIADFGMASLQVDSLLETSCGSPHYACPEVIRGEKYDGRKADVWSCGVILFALLV 128
QY 212 GALPFPDDNLRQLLEKVKRGVFMHPHFIPDQCOSLGRGMIEVDAAARLTLLEHTQKHIWYI 271
DB 129 GALPFPDDNLRQLLEKVKRGVFMHPHFIPDQCOSLGRGMIEVDAAARLTLLEHTQKHIWYI 188
QY 272 GKNPEPEQPIPRKVQIRSLPSLEIDDPVLDMSHSLGCFDRNKLQDLSSEENQEK 331
DB 189 GKNPEPEQPIPRKVQIRSLPSLEIDDPVLDMSHSLGCFDRNKLQDLSSEENQEK 248
QY 332 MIYFLLLDRKERYPSQDEDLPRNEIDPPRKVDSFMLNHRHGRKRRPERKSMELSVTDG 391
DB 249 MIYFLLLDRKERYPSQDEDLPRNEIDPPRKVDSFMLNHRHGRKRRPERKSMELSVTDG 308
QY 392 GSPVPARRAIEMAHQQRSSISGASSGLSTSLSPSRVTPHPSRGSPPLTPKGTVPVHT 451
DB 309 GSPVPARRAIEMAHQQRSSISGASSGLSTSLSPSRVTPHPSRGSPPLTPKGTVPVHT 368

QY 452 PKSPAGTNPPTPPSPSVGGVFWARLNSIKNSFLGSPFRHRRKLVQVTPPEMSNLTP 511
DB 369 PKSPAGTNPPTPPSPSVGGVFWARLNSIKNSFLGSPFRHRRKLVQVTPPEMSNLTP 428
QY 512 SSPSLAKSWFGNFIISLEKEEQIFVVIKDKPLSIKADIVHAFSLSPSLSHSVISQTSFR 571
DB 429 SSPSLAKSWFGNFIISLEKEEQIFVVIKDKPLSIKADIVHAFSLSPSLSHSVISQTSFR 488
QY 572 AEYKATGGPAVFOKPVKFOVDITTEGGEAKENGIIYSVTFTLLSGPSRRFRKRVVETIOA 631
DB 489 AEYKATGGPAVFOKPVKFOVDITTEGGEAKENGIIYSVTFTLLSGPSRRFRKRVVETIOA 548
QY 632 QLLSTHDPAPAAHLSDTTNCMEMMTGRLSKCGIIPKS 668
DB 549 QLLSTHDPAPAAHLSDTTNCMEMMTGRLSKCGIIPKS 585
RESULT 14
US-10-195-071-4
; Sequence 4, Application US/10195071
; Publication No. US20030096271A1
; GENERAL INFORMATION:
; APPLICANT: Origene Technologies
; TITLE OF INVENTION: Full-Length Serine Protein Kinase in Brain and Pancreas
; FILE REFERENCE: 16U 101 C1
; CURRENT APPLICATION NUMBER: US/10/195,071
; CURRENT FILING DATE: 2002-07-15
; PRIOR APPLICATION NUMBER: US 09/930,181
; PRIOR FILING DATE: 2001-08-16
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: Patent in version 3.1
; SEQ ID NO 4
; LENGTH: 585
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-195-071-4
Query Match 86.6%; Score 3028; DB 15; Length 585;
Best Local Similarity 100.0%; Pred. No. 5.5e-188;
Matches 577; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 92 YLVLEHVS GGELFDYLVKKGRLTPKEARKFRQIISALDFCHSHSICHRLDKPENLLIDE 151
DB 9 YLVLEHVS GGELFDYLVKKGRLTPKEARKFRQIISALDFCHSHSICHRLDKPENLLIDE 68
QY 152 KNNIRIADFGWASLQVGSLLTSCGSPHYACPEVIRGEKYDGRKADVWSCGVILFALLV 211
DB 69 KNNIRIADFGWASLQVGSLLTSCGSPHYACPEVIRGEKYDGRKADVWSCGVILFALLV 128
QY 212 GALPDDDDNLRLQLLEKVKRGVFMHPHFIPDQCQLLRGMI EYDAARLITLHIIQKHIWYI 271
DB 129 GALPDDDDNLRLQLLEKVKRGVFMHPHFIPDQCQLLRGMI EYDAARLITLHIIQKHIWYI 188
QY 272 GGNKEPEEQIPRKVQIRSLPSLEDIDPDVLDMSHSLGCFRDRNKLQDILLSEENOEK 331
DB 189 GGNKEPEEQIPRKVQIRSLPSLEDIDPDVLDMSHSLGCFRDRNKLQDILLSEENOEK 248
QY 332 MIYFLLDRKERYPSQEDDLPPRNEIDPPRKRVDS PMLNRHGRKRRPERKSMEVLSVTDG 391
DB 249 MIYFLLDRKERYPSQEDDLPPRNEIDPPRKRVDS PMLNRHGRKRRPERKSMEVLSVTDG 308
QY 392 GSPVPARRAIEAQAQGRSRSISGASGLSTSPVTPHPSPRGSPPLTPKGTPTVHT 451
DB 309 GSPVPARRAIEAQAQGRSRSISGASGLSTSPVTPHPSPRGSPPLTPKGTPTVHT 368
QY 452 PKSPAGTNPPTPPSPSVGGVFWARLNSIKNSFLGSPFRHRRKLVQVTPPEMSNLTP 511
DB 369 PKSPAGTNPPTPPSPSVGGVFWARLNSIKNSFLGSPFRHRRKLVQVTPPEMSNLTP 428
QY 512 SSPSLAKSWFGNFIISLEKEEQIFVVIKDKPLSIKADIVHAFSLSPSLSHSVISQTSFR 571
DB 429 SSPSLAKSWFGNFIISLEKEEQIFVVIKDKPLSIKADIVHAFSLSPSLSHSVISQTSFR 488

QY 572 AEYKATGGPAVFOKPVKFOVDITTEGGEAKENGIIYSVTFTLLSGPSRRFRKRVVETIOA 631
DB 489 AEYKATGGPAVFOKPVKFOVDITTEGGEAKENGIIYSVTFTLLSGPSRRFRKRVVETIOA 548
QY 632 QLLSTHDPAPAAHLSDTTNCMEMMTGRLSKCGIIPKS 668
DB 549 QLLSTHDPAPAAHLSDTTNCMEMMTGRLSKCGIIPKS 585
RESULT 15
US-10-195-072-17
; Sequence 17, Application US/10195072
; Publication No. US20030092036A1
; GENERAL INFORMATION:
; APPLICANT: Origene Technologies
; TITLE OF INVENTION: Full-Length Serine Protein Kinase in Brain and Pancreas
; FILE REFERENCE: 16U 101 C2
; CURRENT APPLICATION NUMBER: US/10/195,072
; CURRENT FILING DATE: 2002-07-15
; PRIOR APPLICATION NUMBER: US 09/930,181
; PRIOR FILING DATE: 2001-08-16
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: Patent in version 3.1
; SEQ ID NO 17
; LENGTH: 603
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-195-072-17
Query Match 86.2%; Score 3015; DB 15; Length 603;
Best Local Similarity 99.7%; Pred. No. 3.9e-187;
Matches 574; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
QY 72 LIEHPHVLKLDVYENKYLVLVLEHVS GGELFDYLVKKGRLTPKEARKFRQIISALDF 131
DB 1 LIEHPHVLKLDVYENKYLVLVLEHVS GGELFDYLVKKGRLTPKEARKFRQIISALDF 60
QY 132 CHSHSICHRLDKPENLLDEKNNIRIADFGWASLQVGSLLTSCGSPHYACPEVIRGEK 191
DB 61 CHSHSICHRLDKPENLLDEKNNIRIADFGWASLQVGSLLTSCGSPHYACPEVIRGEK 120
QY 192 YDGRKADVWSCGVILFALLV GALLPDDDDNLRLQLLEKVKRGVFMHPHFIPDQCQLLRGMI 251
DB 121 YDGRKADVWSCGVILFALLV GALLPDDDDNLRLQLLEKVKRGVFMHPHFIPDQCQLLRGMI 180
QY 252 EYDAARLITLHIIQKHIWYI GGNKEPEEQIPRKVQIRSLPSLEDIDPDVLDMSHSLG 311
DB 181 EYDAARLITLHIIQKHIWYI GGNKEPEEQIPRKVQIRSLPSLEDIDPDVLDMSHSLG 240
QY 312 FRDRNKLQDILLSEENOEKMIYFLLDRKERYPSQEDDLPPRNEIDPPRKRVDS PMLN 371
DB 241 FRDRNKLQDILLSEENOEKMIYFLLDRKERYPSQEDDLPPRNEIDPPRKRVDS PMLN 300
QY 372 RHGKRRPERKSMEVLSVTDG GSPVPARRAIEAQAQGRSRSISGASGLSTSPVTPHPS 431
DB 301 RHGKRRPERKSMEVLSVTDG GSPVPARRAIEAQAQGRSRSISGASGLSTSPVTPHPS 360
QY 432 PHPSPRGSPPLTPKGTPTVHTPKSPAGTNPPTPPSPSVGGVFWARLNSIKNSFLGSP 491
DB 361 PHPSPRGSPPLTPKGTPTVHTPKSPAGTNPPTPPSPSVGGVFWARLNSIKNSFLGSP 420
QY 492 FHRKLVQVTPPEMSNLTPPESSPELAKSWFGNFIISLEKEEQIFVVIKDKPLSIKADIV 551
DB 421 FHRKLVQVTPPEMSNLTPPESSPELAKSWFGNFIISLEKEEQIFVVIKDKPLSIKADIV 480
QY 552 HAFSLSPSLSHSVISQTSFR AEYKATGGPAVFOKPVKFOVDITTEGGEAKENGIIYSVT 611
DB 481 HAFSLSPSLSHSVISQTSFR AEYKATGGPAVFOKPVKFOVDITTEGGEAKENGIIYSVT 540
QY 612 FTLLSGPSRRFRKRVVETIOAQLLSTHDPAPAAHLSD 647
DB 541 FTLLSGPSRRFRKRVVETIOAQLLSTHDPAPAAHLSD 576

Search completed: November 26, 2003, 12:11:01
Job time : 29.2696 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 26, 2003, 11:49:36 ; Search time 40.3103 Seconds
(without alignments)
2630.326 Million cell updates/sec

Title: US-10-054-579-2

Perfect score: 3497

Sequence: 1 MTSTGKGGGAHQAVGVYR.....TNCMEMTGRLSKGLIPKS 668

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1107863 seqs, 158726573 residues

Total number of hits satisfying chosen parameters: 1107863

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A_Geneseq_19Jun03.*

```
1: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1980.DAT.*
2: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1981.DAT.*
3: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1982.DAT.*
4: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1983.DAT.*
5: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1984.DAT.*
6: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1985.DAT.*
7: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1986.DAT.*
8: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1987.DAT.*
9: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1988.DAT.*
10: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1989.DAT.*
11: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1990.DAT.*
12: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1991.DAT.*
13: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1992.DAT.*
14: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1993.DAT.*
15: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1994.DAT.*
16: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1995.DAT.*
17: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1996.DAT.*
18: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1997.DAT.*
19: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1998.DAT.*
20: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1999.DAT.*
21: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA2000.DAT.*
22: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.*
23: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA2002.DAT.*
24: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA2003.DAT.*
```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	3497	100.0	668	23	Human kinase NHP #
2	3497	100.0	668	24	Human serine prote
3	3423	97.9	664	23	Human PKIN-18 prot
4	3374	96.5	674	23	Human protein kina
5	3188	91.2	608	23	Human kinase NHP #
6	3028	86.6	585	24	Human serine prote
7	2799	80.0	778	24	Human kinase #1.
8	2785.5	79.7	794	22	Human protein kina
9	2785.5	79.7	794	23	Human kinase PKIN-

10	2768	79.2	754	24	AA016604	Human cell cycle-r
11	2545.5	72.8	703	24	ABB98745	Human kinase #3.
12	2386.5	68.2	762	24	ABB98744	Human kinase #2.
13	1974	56.4	851	22	ABG62061	Drosophila melanog
14	1275.5	36.5	301	22	ABBI1052	Human HrPOPK-1 hom
15	872.5	24.9	1349	22	AA070854	C albicans apoptos
16	797	22.8	768	22	AA070854	Human protein SEQ
17	796.5	22.8	745	24	AAE33550	Human microtubule
18	796.5	22.8	796	24	ABU11830	Human MDDT polyep
19	794.5	22.7	691	23	AAE19050	Human PAR-1B alpha
20	794.5	22.7	724	23	AAE19051	Human PAR-1B beta
21	793.5	22.7	724	24	ABU08527	Human serine/threo
22	790.5	22.6	514	21	AA034420	Soybean putative c
23	790.5	22.6	514	24	ABR40714	Glycine max oil tr
24	790.5	22.6	722	17	AA098227	Rat neuronal prote
25	788	22.5	769	23	ABP62966	Human polypeptide
26	784.5	22.4	514	24	ABR40816	Glycine max oil tr
27	784.5	22.4	515	21	AB034422	Soybean putative c
28	784.5	22.4	515	24	ABR40716	Glycine max oil tr
29	784	22.4	505	24	ABR40810	Oryza sativa oil t
30	781	22.3	744	24	ABG73795	Murine MARK protei
31	779.5	22.3	761	24	ABP96068	Human protein kina
32	779.5	22.3	752	22	AAE11782	Human kinase (PKIN
33	779.5	22.3	752	22	AA093356	Human polypeptide,
34	779.5	22.3	752	23	AB044433	Human neuronal ser
35	779.5	22.3	752	24	ABG73794	Human MARK protein
36	778.5	22.3	688	24	AAE33554	Human microtubule
37	778.5	22.3	688	24	AAE33555	Human microtubule
38	777.5	22.2	795	23	AAE19052	Human PAR-1C prote
39	777.5	22.2	795	23	AAE16258	Human kinase PKIN-
40	776.5	22.2	795	24	AAE33551	Human microtubule
41	776.5	22.2	804	22	AA079333	Human protein SEQ
42	775.5	22.2	688	23	AB044434	Human neuronal ser
43	774.5	22.1	579	21	AA034417	Corn putative carb
44	774.5	22.1	579	24	ABR40709	Zea mays oil trait
45	774.5	22.1	823	21	AA090879	Human keratinocyte

ALIGNMENTS

RESULT 1

ABB09557
ID ABB09557 standard; Protein; 668 AA.

XX ABB09557;

AC ABB09557;

DT 21-OCT-2002 (first entry)

XX Human kinase NHP #1, SEQ ID NO:2.

DE Human; kinase; NHP; chromosome 11; signal transduction; cytoplasmic;
KW nuclear; brain; pituitary; hypothalamus; adipose; cerebellum;
KW adrenal gland; foetal lung; foetal brain; cancer; drug screening;
KW mental disorder; diagnostic reagent; clinical trial monitoring;
KW nutraceutical; mutation detection; gene expression analysis;
KW chromosome mapping; transgenic animal; cytostatic; gene therapy;
KW enzyme.

XX Homo sapiens.

XX WO200259287-A2.

XX 01-AUG-2002.

XX 22-JAN-2002; 2002WO-US01818.

XX 23-JAN-2001; 2001US-263378P.

XX (LEXI-) LEXICON GENETICS INC.

XX Turner CA, Mathur B;

XX

XX	SQ	Sequence	668 AA;
		Query Match	100.0%; Score 3497; DB 24; Length 668;
		Best Local Similarity	100.0%; Pred. No. 1.6e-249; Indels 0; Gaps 0;
		Matches 668; Conservative	0; Mismatches 0; Indels 0; Gaps 0;
QY	1	MTSTGKGGAQAQVGVPRLEKTLGKGTGLVGLVGHCVTCQKVAIKIVNREKLSVSL	60
Db	1	MTSTGKGGAQAQVGVPRLEKTLGKGTGLVGLVGHCVTCQKVAIKIVNREKLSVSL	60
QY	61	MKVEREIAILKLIIEHPHVLKLDVYENKKYLYLVLEHVS GGELFDYLVKKGRLTPKEARK	120
Db	61	MKVEREIAILKLIIEHPHVLKLDVYENKKYLYLVLEHVS GGELFDYLVKKGRLTPKEARK	120
QY	121	FFRQIISALDFCHSHSICHRDLKPENLLDEKNRIADFGWASLQVGSLLTSCGSPH	180
Db	121	FFRQIISALDFCHSHSICHRDLKPENLLDEKNRIADFGWASLQVGSLLTSCGSPH	180
QY	181	YACPEVIRGEKYDGRKADVWSCGVILFALLVGLPDDNLRQLLEKVKRGVFMHPHIP	240
Db	181	YACPEVIRGEKYDGRKADVWSCGVILFALLVGLPDDNLRQLLEKVKRGVFMHPHIP	240
QY	241	PCQSLLRGMIIEVDAAARLTLEHIQKHIWYIGGKNEPEQPIPRKVQIRSLPSLEDIDP	300
Db	241	PCQSLLRGMIIEVDAAARLTLEHIQKHIWYIGGKNEPEQPIPRKVQIRSLPSLEDIDP	300
QY	301	DVLDSMHSGLCFRDRNKLQDLLEENOEKMIYFLLDRKERYPSQEDDLPPENEIDP	360
Db	301	DVLDSMHSGLCFRDRNKLQDLLEENOEKMIYFLLDRKERYPSQEDDLPPENEIDP	360
QY	361	PKRVDSPLNLRHGRKRRPERKSMELSVTDGGSVPARRAIEAQAQGRSRSISGASSGL	420
Db	361	PKRVDSPLNLRHGRKRRPERKSMELSVTDGGSVPARRAIEAQAQGRSRSISGASSGL	420
QY	421	STSPSSPRVTHPSPRGSLTPKGTVPHTPKESPAGTNPPTPSSVSGVPWRALN	480
Db	421	STSPSSPRVTHPSPRGSLTPKGTVPHTPKESPAGTNPPTPSSVSGVPWRALN	480
QY	481	SIKNSFLGSPRRHRLQVPTPEEKNLTPESSPELAKSWFGNFISLEKEQIFVWIKD	540
Db	481	SIKNSFLGSPRRHRLQVPTPEEKNLTPESSPELAKSWFGNFISLEKEQIFVWIKD	540
QY	541	KPLSSIKADIVHAFISPSLSHSVISQTSFRAEYKATGPAVFKPVPQVDITTEGGE	600
Db	541	KPLSSIKADIVHAFISPSLSHSVISQTSFRAEYKATGPAVFKPVPQVDITTEGGE	600
QY	601	AQKNGIYSVTTLTSGSPRRKRVRVETIQALLSTHDPAAQHLSDTTCNEMMTGRLS	660
Db	601	AQKNGIYSVTTLTSGSPRRKRVRVETIQALLSTHDPAAQHLSDTTCNEMMTGRLS	660
QY	661	KCGIIPKS 668	
Db	661	KCGIIPKS 668	
RESULT 3			
ID	AAE21723		
XX	AAE21723	standard; Protein; 664 AA.	
AC	AAE21723;		
XX	AAE21723;		
DT	16-JUL-2002	(first entry)	
XX	Human	PKIN-18 protein.	
XX	Human	PKIN-18 protein; immune system disorder; anaemia;	
KW	acquired immune deficiency syndrome; thymic hypoplasia; Crohn's disease;		
KW	asthma; neurological disorder; epilepsy; Charcot-Marie-Tooth disease;		
KW	AIDS; seizures; cell proliferative disorder; cancer; adenocarcinoma;		
KW	leukaemia; lymphoma; melanoma; myeloma; sarcoma; developmental disorder;		
KW	Down's syndrome; gene therapy; protein therapy; cytostatic.		

OS	Homo sapiens.
XX	Key
PH	Domain
FT	Location/Qualifiers
FT	15..266
FT	/note= "Eukaryotic protein kinase domain"
FT	16..257
FT	/note= "Protein kinase domain"
FT	17..257
FT	/note= "Protein kinase domain"
FT	36..256
FT	/note= "Protein kinase domain"
XX	WO200218557-A2.
PN	07-MAR-2002.
XX	31-AUG-2001; 2001WO-US27219.
XX	31-AUG-2000; 2000US-229873P.
PR	08-SEP-2000; 2000US-231357P.
PR	14-SEP-2000; 2000US-232654P.
PR	22-SEP-2000; 2000US-234902P.
PR	29-SEP-2000; 2000US-236493P.
PR	06-OCT-2000; 2000US-238389P.
PR	13-OCT-2000; 2000US-240542P.
XX	(INCY-) INCYTE GENOMICS INC.
XX	Bandman O, Nguyen DB, Walia NK, Hafalia AJA, Yao MG, Gandhi AR;
PI	Gururajan R, Ding L, Patterson C, Yue H, Baughn MR, Tribouley CM;
PI	Thornston M, Elliott VS, Lu Y, Ison CH, Au-Young J, Tang YF;
PI	Azimzai Y, Burrill JD, Marcus GA, Zingler KA, Lu DAM, Lal PG;
PI	Ramkumar J, Warren BA, Kearney L, Policky JL, Thangavelu K;
PI	Burford N;
XX	WPI; 2002-329769/36.
DR	N-PSDB; AAD34315.
XX	New human kinases, useful for diagnosing, treating or preventing immune
PT	system disorders (e.g. Crohn's disease), neurological disorders (e.g.
PT	epilepsy), or cell proliferative disorders (e.g. cancers such as
PT	leukemia or lymphoma)
XX	Claim 73; Page 181-183; 218pp; English.
PS	The present invention relates to human kinases (PKIN) and polynucleotides
XX	encoding such proteins. PKIN sequences of the invention are useful for
CC	diagnosing, treating or preventing disorders associated with aberrant
CC	expression of PKIN, particularly immune system disorders (e.g. acquired
CC	immune deficiency syndrome (AIDS), thymic hypoplasia, Crohn's disease,
CC	anaemia, asthma), neurological disorders (e.g. epilepsy, Charcot-Marie-
CC	Tooth disease or seizures), cell proliferative disorders (e.g. cancers
CC	such as adenocarcinoma, leukaemia, lymphoma, melanoma, myeloma, sarcoma),
CC	and developmental disorders (e.g. Down's syndrome). They are also used
CC	in gene therapy and protein therapy. The present sequence is human
CC	PKIN-18 protein.
XX	Sequence 664 AA;
SQ	Query Match 97.9%; Score 3423; DB 23; Length 664;
	Best Local Similarity 100.0%; Pred. No. 4.5e-244;
	Matches 654; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY	15 YVGPYRLEKTLGKGTGLVGLVGHCVTCQKVAIKIVNREKLSVLMKVEREIAILKILIE 74
Db	11 YVGPYRLEKTLGKGTGLVGLVGHCVTCQKVAIKIVNREKLSVLMKVEREIAILKILIE 70
QY	75 HPHVLKLDVYENKKYLYLVLEHVS GGELFDYLVKKGRLTPKEARKFFRQIISALDFCHS 134
Db	71 HPHVLKLDVYENKKYLYLVLEHVS GGELFDYLVKKGRLTPKEARKFFRQIISALDFCHS 130
QY	135 HSIChRLDKPENLLDEKNRIADFGWASLQVGSLLTSCGSPHYACPEVIRGEKYDG 194

Db 131 HSIHRDLKPNLLDKNRIRADFGWASLQVGSLLLETSCGSPHYACPEVIRGEKYDG 190
QY 195 RKADVWSCGVILFALLVGCALPFDDNLRQLLEKVKRGVFMHPHIPPDCQSLLRGMIEVD 254
Db 191 RKADVWSCGVILFALLVGCALPFDDNLRQLLEKVKRGVFMHPHIPPDCQSLLRGMIEVD 250
QY 255 AARRLTLEHIQKHIWYIGGKNEPEQPIPRKVOIRSLPSLEDDPDVLDMSHSLGCFRD 314
Db 251 AARRLTLEHIQKHIWYIGGKNEPEQPIPRKVOIRSLPSLEDDPDVLDMSHSLGCFRD 310
QY 315 RNKLLQDLLSEENQEKMIYFLLDRKERYPSQDEDDLPNNEIDPPRKRVDSPMLNRHG 374
Db 311 RNKLLQDLLSEENQEKMIYFLLDRKERYPSQDEDDLPNNEIDPPRKRVDSPMLNRHG 370
QY 375 KRRPERKMEVLSVTDGSPVPARRAIEAQAQHGORSISGASSGLSTSPSSPRVTPHP 434
Db 371 KRRPERKMEVLSVTDGSPVPARRAIEAQAQHGORSISGASSGLSTSPSSPRVTPHP 430
QY 435 SPRGSPPLTPKGTPTVHTPKESPAGTNPPTPSSSVGGVPPWRARLNSIKNSFLGSPRFR 494
Db 431 SPRGSPPLTPKGTPTVHTPKESPAGTNPPTPSSSVGGVPPWRARLNSIKNSFLGSPRFR 490
QY 495 RKLQVTPTEENSLTPSSPELAKKSWFGNFIISLEKQEIFVWIKDKPLSSIKADIVHAF 554
Db 491 RKLQVTPTEENSLTPSSPELAKKSWFGNFIISLEKQEIFVWIKDKPLSSIKADIVHAF 550
QY 555 LSIPSLSHSVISQTSFRAEYKATGGPAVFOKPVKQVDITYTEGGAOKENGIYSVTFTL 614
Db 551 LSIPSLSHSVISQTSFRAEYKATGGPAVFOKPVKQVDITYTEGGAOKENGIYSVTFTL 610
QY 615 LSGPSRRFRVVTETIQALLSTHDPAAQHLSDTTNCMEMMTGRLSKCGIIPKS 668
Db 611 LSGPSRRFRVVTETIQALLSTHDPAAQHLSDTTNCMEMMTGRLSKCGIIPKS 664

RESULT 4
AA047830
ID AA047830 standard; Protein; 674 AA.
AC AA047830;
XX
DT 19-FEB-2002 (first entry)
XX
DE Human protein kinase 2246 SEQ ID NO 2.
XX
KW Human; protein kinase 2246; cytostatic; immunomodulator; carcinoma;
KW anti-inflammatory; analgesic; cardiovascular; cancer; sarcoma;
KW cellular proliferation disorder; cellular differentiation disorder;
KW metastatic; haematopoietic disorder; leukaemia; immune disorder;
KW inflammatory disorder; arthritis; autoimmune disease; diabetes mellitus;
KW psoriasis; Crohn's disease; cardiovascular disease; virus; pain;
KW gene therapy.
XX
OS Homo sapiens.
XX
PN WO200181588-A2.
XX
PD 01-NOV-2001.
XX
PF 25-APR-2001; 2001WO-US13784.
XX
PR 25-APR-2000; 2000US-199391P.
XX
PA (MILL-) MILLENNIUM PHARM INC.
XX
PI Meyers R;
XX
DR WPI; 2002-049281/06.
XX
DR N-PSDB; ABA02994, ABA02995.
XX
PT New protein kinase nucleic acid and polypeptide molecules, designated
PT 2246, useful for diagnosing, preventing or treating cancer or a
PT cellular proliferation/differentiation disorders, e.g. carcinoma,

PT sarcoma or leukaemias -
XX Claim 4; Fig 1; 11pp; English.
XX
CC The invention relates to the human protein kinase 2246 gene and the
CC the isolated encoded polypeptide with cytostatic, immunomodulator,
CC anti-inflammatory, analgesic and cardiovascular activity. The 2246
CC nucleic acid and polypeptide are useful for diagnosing, preventing or
CC treating a subject having cancer or a cellular proliferation and/or
CC differentiation disorder or at risk of developing cancer or a cellular
CC proliferation and/or differentiation disorder. In particular, the
CC disorder includes carcinoma, sarcoma, metastatic or haematopoietic
CC disorders (e.g. leukaemias) or cancers of the lung, breast, thyroid, head
CC neck, prostate or genito-urinary tract. The 2246 nucleic acid and
CC polypeptide are also useful for treating immune disorders, e.g.
CC inflammatory (e.g. respiratory inflammation or arthritis), autoimmune
CC disease (e.g. diabetes mellitus, psoriasis, Wegener's granulomatosis,
CC Crohn's disease or Grave's disease), for treating cardiovascular
CC diseases, endothelial cell disorder, viral diseases or pain. The nucleic
CC acid and polypeptide are also useful for evaluating the efficacy of a
CC treatment of cancer or a cellular proliferation and/or differentiation
CC disorder. The nucleic acid is useful for gene therapy. The present
CC sequence is that of the 2246 protein kinase.
XX
SQ Sequence 674 AA;
Query Match 96.5%; Score 3374; DB 23; Length 674;
Best Local Similarity 99.7%; Pred. No. 1.9e-240;
Matches 645; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
QY 1 MTSTGDKGGAQAQYVGYPRLEKTLGKGTGLVKLGVCVTCOKVAIKVNRKLSVSL 60
Db 1 MTSTGDKGGAQAQYVGYPRLEKTLGKGTGLVKLGVCVTCOKVAIKVNRKLSVSL 60
QY 61 MKVERETAILKLEHPLVXLHDVYENKYLVLVLEHVSGGELFDYLVKGRLLTPKEARK 120
Db 61 MKVERETAILKLEHPLVXLHDVYENKYLVLVLEHVSGGELFDYLVKGRLLTPKEARK 120
QY 121 FFRQIISALDFCHSHSICHRLDKPENLLDEKKNIRIADFGMASLQVDSLSLSCSPH 180
Db 121 FFRQIISALDFCHSHSICHRLDKPENLLDEKKNIRIADFGMASLQVDSLSLSCSPH 180
QY 181 YACPEVIRGEKYDGRKADVWSCGVILFALLVGCALPFDDNLRQLLEKVKRGVFMHPHIF 240
Db 181 YACPEVIRGEKYDGRKADVWSCGVILFALLVGCALPFDDNLRQLLEKVKRGVFMHPHIF 240
QY 241 PDCQSLLRGMIEVDAAARLTLEHIQKHIWYIGGKNEPEQPIPRKVOIRSLPSLEDD 300
Db 241 PDCQSLLRGMIEVDAAARLTLEHIQKHIWYIGGKNEPEQPIPRKVOIRSLPSLEDD 300
QY 301 DVLDSMHSLSGCFRDRNKLQDLLSEENQEKMIYFLLDRKERYPSQDEDDLPNNEIDP 360
Db 301 DVLDSMHSLSGCFRDRNKLQDLLSEENQEKMIYFLLDRKERYPSQDEDDLPNNEIDP 360
QY 361 PRKRVDSPLNLRHGRKRRPERKSMESVLSVTDGSPVPARRAIEAQAQHGORSISGASSGL 420
Db 361 PRKRVDSPLNLRHGRKRRPERKSMESVLSVTDGSPVPARRAIEAQAQHGORSISGASSGL 420
QY 421 STSPSSPRVTPHPSPRGSPPLTPKGTPTVHTPKESPAGTNPPTPSSSVGGVPPWRARLN 480
Db 421 STSPSSPRVTPHPSPRGSPPLTPKGTPTVHTPKESPAGTNPPTPSSSVGGVPPWRARLN 480
QY 481 SIKNSFLGSPRFRHGRKRRKQVTPTEENSLTPSSPELAKKSWFGNFIISLEKQEIFVWIKD 540
Db 481 SIKNSFLGSPRFRHGRKRRKQVTPTEENSLTPSSPELAKKSWFGNFIISLEKQEIFVWIKD 540
QY 541 KPLSSIKADIVHAFSLSPSLSHSVISQTSFRAEYKATGGPAVFOKPVKQVDITYTEGGE 600
Db 541 KPLSSIKADIVHAFSLSPSLSHSVISQTSFRAEYKATGGPAVFOKPVKQVDITYTEGGE 600
QY 601 AQKENGISYVTFLLSGPSRRFRKRVVETIQALLSTHDPAAQHLSD 647
Db 601 AQKENGISYVTFLLSGPSRRFRKRVVETIQALLSTHDPAAQHLSD 647

RESULT 5
ABB09558
ID ABB09558 standard; Protein; 608 AA.
XX AC ABB09558;
XX DT 21-OCT-2002 (first entry)
XX DE Human kinase NHP #2, SEQ ID NO:4.
XX KW Human; kinase; NHP; chromosome 11; signal transduction; cytoplasmic;
KW nuclear; brain; pituitary; hypothalamus; adipose; cerebellum;
KW adrenal gland; foetal lung; foetal brain; cancer; drug screening;
KW mental disorder; diagnostic reagent; clinical trial monitoring; cosmetic;
KW nutraceutical; mutation detection; gene expression analysis;
KW chromosome mapping; transgenic animal; cytostatic; gene therapy;
KW enzyme.
XX OS Homo sapiens.
XX PN W0200259287-A2.
XX PD 01-AUG-2002.
XX PF 22-JAN-2002; 2002WO-US01818.
XX PR 23-JAN-2001; 2001US-263378P.
XX PA (LEXI-) LEXICON GENETICS INC.
XX PI Turner CA, Mathur B;
XX WPI; 2002-599780/64.
XX DR N-ESDB; ABQ77626.
XX PT Novel polynucleotide encoding human proteins sharing sequence
PT similarity with animal kinases, useful for drug screening, diagnosis,
PT in gene therapy of disorders and diseases e.g. cancer -
XX Claim 4; Page 39-40; 40pp; English.
XX The invention relates to novel human kinases (ABB09557, AAB09558),
XX referred to as NHPs, and to nucleic acids encoding them (ABQ77625,
XX ABQ77626). The NHPs have structural similarity with animal kinases
XX including serine-threonine kinases, carbon catabolite repressing kinases,
XX calcium/calmodulin-dependent protein kinases, and CAMP-dependent protein
XX kinases. The NHPs do not contain consensus signal sequences, indicating
XX that they may be cytoplasmic or nuclear proteins, and are thought to be
XX involved in signal transduction. Polynucleotides encoding NHPs were
XX obtained using human brain, pituitary, hypothalamus, adipose, cerebellum,
XX adrenal gland, foetal lung, and foetal brain cDNAs and primers derived
XX from human genomic DNA. The gene encoding the NHPs is located on
XX chromosome 11. NHP nucleotides and proteins are useful for treating
XX mental, biological or medical disorders including cancer, and for
XX screening compounds useful in the treatment of such conditions. They can
XX also be used as diagnostic reagents, in clinical trial monitoring and in
XX cosmetic and nutraceutical applications. NHP nucleotides can additionally
XX be used in the detection of disease-associated mutations, in the analysis
XX of gene expression, for mapping chromosome 11, for the recombinant
XX expression of NHPs, to generate transgenic animals, in gene therapy,
XX and as part of ribozyme and/or triple helix sequences useful in the
XX modulation of NHP gene expression. The present sequence represents a
XX 608 amino acid NHP (#2).

Query Match 91.2%; Score 3188; DB 23; Length 608;
Best Local Similarity 100.0%; Pred.No. 9e-227;
Matches 608; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
61 MKVEREIALKLIIEHPVLYLHDVYENKYLVLVLEHVS GGE LFDYLVKKGR LTPKEARK 120

Db 1 MKVEREIALKLIIEHPVLYLHDVYENKYLVLVLEHVS GGE LFDYLVKKGR LTPKEARK 60
Qy 121 FFRQIIISALDFCHSHSICHRLDKPENLLLEKKNIRIADFGMASLQVDSLLTSCGSPH 180
Db 61 FFRQIIISALDFCHSHSICHRLDKPENLLLEKKNIRIADFGMASLQVDSLLTSCGSPH 120
Qy 181 YACPEVIRGEKDYGRKADVWSCGVILFALLVGLALPEDDNLROLLEKVKGVFHMHPFIP 240
Db 121 YACPEVIRGEKDYGRKADVWSCGVILFALLVGLALPEDDNLROLLEKVKGVFHMHPFIP 180
Qy 241 PDCQSLLRGMIEVDAARRLTLEHIQKHIWYIGGKNEPEPEQIPRKVQIRSLPSLEIDP 300
Db 181 PDCQSLLRGMIEVDAARRLTLEHIQKHIWYIGGKNEPEPEQIPRKVQIRSLPSLEIDP 240
Qy 301 DVLDMSHSLGCFDRNKLQDLLESEENQEKMIYFLLDKERYPSQEDDLPPRNEIDP 360
Db 241 DVLDMSHSLGCFDRNKLQDLLESEENQEKMIYFLLDKERYPSQEDDLPPRNEIDP 300
Qy 361 PRKRVDSPLMLNRHGKRRPERKSMELSVTDGGSVPARRAIEWAHQGRSRSISGASSGL 420
Db 301 PRKRVDSPLMLNRHGKRRPERKSMELSVTDGGSVPARRAIEWAHQGRSRSISGASSGL 360
Qy 421 STSPLSSPRVTPHPSPRGSPPLTPKGTVPVHTPKESPAGTPNPTPPSSPSVGGVPWRARLN 480
Db 361 STSPLSSPRVTPHPSPRGSPPLTPKGTVPVHTPKESPAGTPNPTPPSSPSVGGVPWRARLN 420
Qy 481 SIKNSFLGSPRFRHRKLOVPTPEEMNLTPESPPELAKSWFONFISLEKEQIFVVIKD 540
Db 421 SIKNSFLGSPRFRHRKLOVPTPEEMNLTPESPPELAKSWFONFISLEKEQIFVVIKD 480
Qy 541 KPLSSSIKADIVHAFLSIPSLSHSVISQTSFRABYKATGGPAVFQKPKFOVDITYTEGGE 600
Db 481 KPLSSSIKADIVHAFLSIPSLSHSVISQTSFRABYKATGGPAVFQKPKFOVDITYTEGGE 540
Qy 601 AQKNGIYSVTFTLLSGSPRRFRKRVVETIIQAQLLSTHDPAAQHLSDTTNCMEMMTGRLS 660
Db 541 AQKNGIYSVTFTLLSGSPRRFRKRVVETIIQAQLLSTHDPAAQHLSDTTNCMEMMTGRLS 600
Qy 661 KCGIIPKS 668
Db 601 KCGIIPKS 608
RESULT 6
ABG72383
ID ABG72383 standard; Protein; 585 AA.
XX AC ABG72383;
XX DT 10-FEB-2003 (first entry)
XX DE Human serine protein kinase KSE336-2.
XX Human; enzyme; chromosome 11p15.5-pter; astrocytoma; meningioma;
XX pancreatic adenocarcinoma; insulin-dependent diabetes mellitus 2;
XX helicoid peripapillary choroidretinal degeneration; brain; pancreas;
XX Beckwith-Wiedemann syndrome; congenital hyperinsulinism; KSE336.
XX OS Homo sapiens.
XX PN US6455292-B1.
XX PD 24-SEP-2002.
XX PF 16-AUG-2001; 2001US-0930181.
XX PR 16-AUG-2001; 2001US-0930181.
XX PA (ORIG-) ORIGENE TECHNOLOGIES INC.
XX PI Shu Y, Fan W, Kovacs KF, Zidanic M, Jay G;
XX

Db 301 GRRVAMSLPSNGELDPDVLSEMASLGCFFDRRLHRELSEENQEKMIYLLDRKER 360
QY 344 YPSQDEDLPPNETDPRKRVDSPLNPHGKRRPERKSMVLSVTD---GGSPVPARRA 400
Db 361 YPSQDQDLPPNDVDPKRVDSPLNPHGKRRPERKSMVLSITDAGGGGSPVPTERRA 420
QY 401 IEMAQHGQSRISGASSGLSTSPSSPR----- 429
Db 421 LEMAHQSQRSSRVSGASTGLSSPLSSPRSPVFSFPEPGAGDEARGGSPSTKTQLPS 480
QY 430 -----VTPHPSPGSPPLTPKG-----TPVHTPKESPAGTNPPTPPSP 468
Db 481 RGRPGGAGEQPPPSARSSTPLPGPGPRSSGGTPLHSPHTPRASPTGPTPPSP 540
QY 469 --SVGVPWRARLNSIKNSFLGSPFRHKLQVPTPEMSNLTPESSPELAKSWFGNFI 526
Db 541 GGGVGAARSLNSIRNSFLGSPFRHKLQVPTAEENSLTPESSPELAKSWFGNFI 600
QY 527 SLEKEEQIFVVIKDKPLSSIKADIHAFSLIPSLSHSVISQTSFRAEYKATGGPAVFOKP 586
Db 601 SLDKEEQIFVLKDKPLSSIKADIHAFSLIPSLSHSVISQTSFRAEYKATGGPAVFOKP 660
QY 587 VKFQVDITVTEGEQAQKE-----NGIYSVTFTLLSGPSRRFRKRVVETIQALLSTHOPP 640
Db 661 VRFQVDISSEGPSPRRDGGGGGIYSVTFTLLSGPSRRFRKRVVETIQALLSTHOPP 720
QY 641 AAQHLSDTTN 650
Db 721 SVQALADEKN 730

RESULT 8

AAU03517

ID AAU03517 standard; Protein; 794 AA.

AC AAU03517;

XX

DT 12-SEP-2001 (first entry)

DE Human protein kinase #17.

XX

KW Human; protein kinase; PTK; STK; cancer; cardiovascular disease;

KW metabolic disorder; immune related disease; neurological disorder;

KW neurodegenerative disorder; inflammatory disorder; infectious disease;

KW reproductive disorder.

XX

OS Homo sapiens.

XX

FN WO200138503-A2.

XX

PD 31-MAY-2001.

XX

PF 22-NOV-2000; 2000WO-US32085.

XX

PR 24-NOV-1999; 99US-0167482.

XX

PA (SUC-) SUGEN INC.

XX

PI Plowman GD, Whyte D, Manning G, Sudareanam S, Martinez R;

PI Flanagan P, Clary D;

XX

DR WPI; 2001-343950/36.

XX

DR N-PSDB; AAS06717.

XX

PT Nucleic acids encoding human kinase polypeptides, useful for preventing

PT diagnosing and/or treating e.g. cancer, immune, cardiovascular and

PT neuronal-associated diseases, and microbial infections -

XX

PS Claim 7; Figure 2; 433pp; English.

XX

CC AAU03501-AAU03557 represent novel human protein kinases #1-57. The

CC novel protein kinases have been identified as members of the tyrosine

or serine/threonine kinase (PTK and STK) families. The polynucleotides encoding protein kinases and the polypeptides may be used in the prevention, diagnosis and treatment of diseases associated with cancers (especially cancers of haematopoietic origin), cardiovascular disease (e.g. atherosclerosis), metabolic disorders (e.g. diabetes), immune related diseases (e.g. rheumatoid arthritis), neurological disorders (e.g. schizophrenia), neurodegenerative disorders (e.g. Parkinson's disease), inflammatory disorders (e.g. asthma), infectious disease (e.g. HIV) and reproductive disorders (e.g. infertility). Additionally, polynucleotides encoding protein kinases may be used for gene therapy and as DNA probes in diagnostic assays. The protein kinase polypeptides may be used as antigens in the production of antibodies against the protein kinases and in assays to identify modulators of protein kinase expression and activity.

XX Sequence 794 AA;

Query Match 79.7%; Score 2785.5; DB 22; Length 794;

Best Local Similarity 76.9%; Pred. No. 6.7e-197;

Matches 542; Conservative 53; Mismatches 45; Indels 65; Gaps 7;

QY 11 QHAYVGPYRLKTLGKQTLGVLKGVHCVTCQKVAIKIVNREKLSBSVLMKVEREAIL 70

Db 42 RHAQYVGPYRLKTLGKQTLGVLKGVHCVTCQKVAIKIVNREKLSBSVLMKVEREAIL 101

QY 71 KLIHHPHVLKLDVYENKLYLVLEHVSGLGELFDYLVKKGRLLTPKEARKEFRQIISALD 130

Db 102 KLIHHPHVLKLDVYENKLYLVLEHVSGLGELFDYLVKKGRLLTPKEARKEFRQIISALD 161

QY 131 FCHSHSICHRDLKPENLLLOKKNIRIADFGMASLQVGDLSLETS CGSPHYACPEVIRGE 190

Db 162 FCHSYISICHRDLKPENLLLOKKNIRIADFGMASLQVGDLSLETS CGSPHYACPEVIRGE 221

QY 191 KYDGRKADVMSCGVILFALLVWALPDDNLRQLLEKVKRGVFMHPIFPDQCSSLRGM 250

Db 222 KYDGRADVMSCGVILFALLVWALPDDNLRQLLEKVKRGVFMHPIFPDQCSSLRGM 281

QY 251 IEVDAARRLTLEHTQKHIWYIGKNEPEP-EQPTP-RKVQIRSLPSLEDIDPDVLDMSHS 308

Db 282 IEVPEPRKLSLEQIQKHPWYIGGHEPDCLEPAGRRVAMRSLPSNGELDPDVLSENAS 341

QY 309 LGCFDRNKLQDLLSEENQEKMIYFLLLDKRYPSQEDDLP PRNEIDPPKRVDS 368

Db 342 LGCFDRDRRLHRELSEENQEKMIYFLLLDKRYPSQEDDLP PRNDVDPKRVDS 401

QY 369 MLNRHGKRRPERKSMVLSVTD---GGSPVPARRAEMAQHGQSRISGASSGLSTSP 425

Db 402 MLSRHGKRRPERKSMVLSITDAGGGSPVTRRALEMAQHGQSRISGASTGLSSPL 461

QY 426 SSPP-----VTPHPSPGSPPLTP 444

Db 462 SSPPSPVFSFPEPGAGDEARGGSPSTKTQLPSRPGGGAGEQPPPSARSSTPLPGP 521

QY 445 KG-----TPVHTPKESPAGTNPPTPPSP--SVGVPWRARLNSIKNSFLGSPR 491

Db 522 PGSPRSSGGTPLHSPHTPRASPTGPTTTPPPSGGVGGAARSLNSIRNSFLGSPR 581

QY 492 FHRRLQVPTPEMSNLTPESSPELAKSWFGNFI SLEKEEQIFVVIKDKPLSSIKADI 551

Db 582 FHRRLQVPTAEENSLTPESSPELAKSWFGNFI SLDKEEQIFVLKDKPLSSIKADI 641

QY 552 HAFSLIPSLSHSVISQTSFRAEYKATGGPAVFOKPVKFQVDITVTEGEQAQKE-----N 605

Db 642 HAFSLIPSLSHSVISQTSFRAEYKATGGSPVFOKVPVQVDISSEGPSPRRDGGGG 701

QY 606 GIYSVTFTLLSGPSRRFRKRVVETIQALLSTHDPAPAAQHLSDDTTN 850

Db 702 GIYSVTFTLLSGPSRRFRKRVVETIQALLSTHDPQSVQALADEKN 746

RESULT 9

AAE16271

AC AAO16604;
XX 08-MAY-2003 (first entry)
XX Human cell cycle-regulatory factor Cdr2.
XX Human; cell cycle-regulatory factor; Cdr2; kinase; proliferative disease;
KW anticancer agent; wound-healing drug.
XX Homo sapiens.
XX WO200299110-A1.
XX 12-DEC-2002.
XX 03-JUN-2002; 2002WO-JP05411.
XX 04-JUN-2001; 2001JP-0168792.
XX (TAIH) TAIHO PHARM CO LTD.
XX (NAKA/) NAKANISHI M.
XX Nakanishi M;
XX WPI; 2003-156857/15.
XX N-PSDB; AAL51989.
XX Cell cycle-regulatory factor Cdr2 with kinase activity and encoded
PT gene, applicable in diagnosis of and screening drugs for proliferative
PT diseases e.g. anticancer agents and wound-healing drugs -
XX Claim 1; Fig 2 A-C; 63pp; Japanese.
XX The invention comprises the amino acid and coding sequence of the human
CC cell cycle-regulatory factor Cdr2 (with kinase activity). The DNA and
CC protein sequences of the invention are useful in diagnosing and screening
CC drugs for proliferative diseases (e.g. anticancer agents and wound-
CC healing drugs). The present amino acid sequence represents the human cell
CC cycle-regulatory factor Cdr2 protein.
XX Sequence 754 AA;
Query Match 79.2%; Score 2768; DB 24; Length 754;
Best Local Similarity 78.9%; Pred. No. 1.2e-195;
Matches 537; Conservative 52; Mismatches 51; Indels 41; Gaps 8;
QY 11 QHAYVGPYRLKLTGKQOTGLVGLGVHCVTCQKVAIKIVNREKLSVLMKVERETAIL 70
DB 26 QHAYVGPYRLKLTGKQOTGLVGLGVHCVTCQKVAIKIVNREKLSVLMKVERETAIL 85
QY 71 KLIHHPVHLKLDVYENKYLVLVLEHVSQGGELFDYLVKGRTPKEARKPFROIISALD 130
DB 86 KLIHHPVHLKLDVYENKYLVLVLEHVSQGGELFDYLVKGRTPKEARKPFROIIVCAG 145
QY 131 FCHSHSICHRDLKPNLLDKNNIRIADFQWASLQVGSLLSTSCGSPHYACPEVIRGE 190
DB 146 LCHSYICHRDLKPNLLDKNNIRIADFQWASLQVGSLLSTSCGSPHYACPEVIRGE 205
QY 191 KYDGRKADVWCGVLLFALLVGLPFDNNLRQLLEKVRGVFHPHPIPPDCOSLLRGM 250
DB 206 KYDGRADWVCGVLLFALLVGLPFDNNLRQLLEKVRGVFHPHPIPPDCOSLLRGM 265
QY 251 IEVDAARELLTLEHIOKHWIYGGKNEPEP-EQIP-RKVQIRSLPSLEDIDPDVLDMSHS 308
DB 266 IEVEPEKLSLEQIKHPYILGGKHEPFCLEPAPGRVAMRSLPSNGELDPDVLESWAS 325
QY 309 LGCFFDRNKLQDLSSEENQKMTYFLLDRKERYPSQEDDELPFRNEIDPPRKRVDSP 368
DB 326 LGCFFDRRLHRELSEENQKMTYFLLDRKERYPSQEDDELPFRNDVDPKRKRVDSP 385
QY 369 MLNRHGKRPKRKSNEVLVSTVD--GSPVPARRAIEAHOQRSRSISGASSGLSTSP 425
DB 386 MLNRHGKRPKRKSNEVLVSTVDGAGGGSFVPTRRALEMAHQSRSSVSGASTGLSSPL 445

QY 426 SSPP-----VTPH-----PSPRGSLP---TPKGPVHTPKESPA 457
DB 446 SSPPSPVFSFSPGAGDEARGGSGTSTKTQTLPGPPGSPRSGGTLHSLHTPTPRASPT 505
QY 458 GTNPPTPPSSP--SVGGVPWRARLNSIKNSFLGSPRHRRLKQVPTPEMNSLTPSSPE 515
DB 506 GTPGTTTPPSPGSGVGAARSLNIRNSFLGSPRHRRLKQVPTPEMNSLTPSSPE 565
QY 516 LAKKSWFGNFISSLEKEQIFVWIKDKPLSSIKADIVHAFSLPSLSHVSISQTSFRAEYK 575
DB 566 LAKKSWFGNFISSLEKEQIFVWIKDKPLSSIKADIVHAFSLPSLSHVSISQTSFRAEYK 625
QY 576 ATGGPAVFOKPVKQVDFQVDTITVTEGGEAQKE-----NGIYSVTFTLSGSPRRFRKRVVETI 629
DB 626 ASGGPSVFQKPVKQVDFQVDTISSEGEPEPSPRRDGGGGGIYSVTFTLSGSPRRFRKRVVETI 685
QY 630 QAQLLSTHDPAAQHLSDTTN 650
DB 686 QAQLLSTHDPQSVQALADEKN 706

RESULT 11

ABB98745
ID ABB98745 standard; Protein; 703 AA.

XX AC ABB98745;

DT 20-JAN-2003 (first entry)

XX DE Human kinase #3.

XX KW Human; kinase; chromosome 19.

XX OS Homo sapiens.

XX PN WO200281670-A1.

XX PD 17-OCT-2002.

XX PF 04-APR-2002; 2002WO-US10786.

XX PR 06-APR-2001; 2001US-282036P.

XX PA (LEXI-) LEXICON GENETICS INC.

XX PI Turner CA, Mathur B, Friddle CJ;

XX DR WPI; 2003-058538/05.

XX DR N-PSDB; ABV74559.

XX PT New human kinase proteins useful for diagnosis, drug screening, and
PT clinical trial monitoring, treatment of disorders and diseases, and
PT cosmetic and nutritional applications -

XX PS Claim 5; Page 44-46; 47pp; English.

XX CC The present sequence is a novel human kinase. The genomic locus encoding
CC the kinase is thought to be on human chromosome 19. The kinase and its
CC coding sequence are useful for diagnosis, drug screening, clinical trial
CC monitoring, treatment of disorders and diseases, and cosmetic and
CC nutritional applications.

XX SQ Sequence 703 AA;

Query Match 72.8%; Score 2545.5; DB 24; Length 703;

Best Local Similarity 75.6%; Pred. No. 3e-179;

Matches 495; Conservative 51; Mismatches 44; Indels 65; Gaps 7;

QY 61 MKVERETAILKLIHHPVHLKLDVYENKYLVLVLEHVSQGGELFDYLVKGRTPKEARK 120

DB 1 MKVERETAILKLIHHPVHLKLDVYENKYLVLVLEHVSQGGELFDYLVKGRTPKEARK 60

121 FFRQIISALDFCHSHSICHRLDKPENLLDKNNIRIADFGMASLQVGDLSLETSCGSPH 180
 61 FFRQIVSALDFCHSYSICHRLDKPENLLDKNNIRIADFGMASLQVGDLSLETSCGSPH 120
 181 YACEVIRGEKYGDRKADWSCGVLFPDNNLRQLLEKVKRGVFMHPHIP 240
 121 YACEVIRGEKYGDRKADWSCGVLFPDNNLRQLLEKVKRGVFMHPHIP 180
 241 PDCQSLRGMIEVDAAARLTLEHIOKIWIYIGGKNEPEP-EQIP-RKVOIRSLPSLEDI 298
 181 PDCQSLRGMIEVEPEKLSLEQKQKWPYLGKHEDPCLEPAPGRVAMRSLPSNGEL 240
 299 DPVDLSMHSLGCFRDRNKLQDLSSEENQEKMIYFLLDKRKERYPSQEDDLPPENEI 358
 241 DPVLESWASLGCFRDRERLHRELRSSEENQEKMIYVLLDKRKERYPSCEDQDLPPNDV 300
 359 DPPKRVDSPLNLRHKKRPERKSMVLSVTD---GGSPVAPRAIEMAHQGRSRISG 415
 301 DPPKRVDSPLNLRHKKRPERKSMVLSITDAGGGGSPVTRDALEMAHQSRSRISG 360
 416 ASSGLSTSPSSPR-----VTPHP 434
 361 ASTGLSSPLSPSPVPSFSPFEGAGDEARGGSPSTKQTLPSRGRGGGAGEQPPPP 420
 435 SPRGSPPTPKG-----TPVHTPKESPAGTNPPTPSSP--SVGGVPMRRLNS 481
 421 SARSTLPSPGPGSPSSSGTPLHSPHTPRASPTGTPTTTPPSFGGCVGGAWRSLNS 480
 482 INKSFSGPRHRRKLQVPTPEENSLTPSSPELAKSWFGNFISLEKEQIFVVLKDK 541
 481 IRNSFLGSPRHRKMQVPTAEENSSLTPESSPELAKSWFGNFISLKEEQIFVLKDK 540
 542 PLSSTKADIVHAFSLPSLSHVSISQTSFRAEYKATGGPAVKFQVDFVITTEGEA 601
 541 PLSSTKADIVHAFSLPSLSHVSISQTSFRAEYKASGGSPVQKPVRFQVDISSEGEPE 600
 602 QKE-----NGIYSVFTLLSGPSRRFRKVVETIOALLSTHDPPAAQHLSDDTN 650
 601 SPRRDGGGGGIYSVFTLLSGPSRRFRKVVETIOALLSTHDQPSVQALADEKN 655

RESULT 12

ABB98744
 ID ABB98744 standard; Protein; 762 AA.

AC ABB98744;

XX 20-JAN-2003 (first entry)

XX Human kinase #2.

XX Human; kinase; chromosome 19.

XX Homo sapiens.

XX WO200281670-A1.

XX 17-OCT-2002.

XX 04-APR-2002; 2002WO-US10786.

XX 06-APR-2001; 2001US-282036P.

XX (LEXI-) LEXICON GENETICS INC.

XX Turner CA, Mathur B, Friddle CJ;

XX WPI; 2003-058538/05.

XX N-PSDB; ABV74558.

XX New human kinase proteins useful for diagnosis, drug screening, and
 PT clinical trial monitoring, treatment of disorders and diseases, and
 PT cosmetic and nutritional applications

XX Claim 5; Page 42-43; 47pp; English.
 XX The present sequence is a novel human kinase. The genomic locus encoding
 CC the kinase is thought to be on human chromosome 19. The kinase and its
 CC coding sequence are useful for diagnosis, drug screening, clinical trial
 CC monitoring, treatment of disorders and diseases, and cosmetic and
 CC nutritional applications.

XX Sequence 762 AA;

Query Match 68.2%; Score 2386.5; DB 24; Length 762;
 Best Local Similarity 74.1%; Pred. No. 1.8e-167;
 Matches 464; Conservative 52; Mismatches 45; Indels 65; Gaps 7;

QY 90 YLVLEHVGSGELFDYLVKGRLLTPKEARKFRQIISALDFCHSHSICHRLDKPENLL 149
 DB 89 FRYLVLEHVGSGELFDYLVKGRLLTPKEARKFRQIISALDFCHSHSICHRLDKPENLL 148
 QY 150 DEKNIRIADFGMASLQVGDLSLETSCGSPHYACPEVIRGEKYDGRKADWSCGVILFAL 209
 DB 149 DEKNIRIADFGMASLQVGDLSLETSCGSPHYACPEVIRGEKYDGRADWSCGVILFAL 208
 QY 210 LVGALPFDNNLRQLLEKVKRGVFMHPHIPDCCQLLRGMIEVDAAARLTLEHIOKHIW 269
 DB 209 LVGALPFDNNLRQLLEKVKRGVFMHPHIPDCCQLLRGMIEVEPEKLSLEQIQHPW 268
 QY 270 YIGGKNEPEP-EQIP-RKVOIRSLPSLEDDIPDVLDSMHSLGCFRDRNKLQDLSSEE 327
 DB 269 YIGGKNEPEP-EQIP-RKVOIRSLPSLEDDIPDVLDSMHSLGCFRDRERLHRELSEEE 328
 QY 328 NOEKMIYVLLDKRKERYPSCEDQDLPPRNDVDPKRVDSPLNLRHKKRPERKSMVLS 387
 DB 329 NOEKMIYVLLDKRKERYPSCEDQDLPPRNDVDPKRVDSPLNLRHKKRPERKSMVLS 388
 QY 388 VTP---GGSPVAPRAIEMAHQGRSRISQTSFRAEYKATGGPAVKFQVDFVITTEGEA 429
 DB 389 ITDAGGGGSPVTRDALEMAHQSRSRISQTSFRAEYKATGGPAVKFQVDFVITTEGEA 448
 QY 430 -----VTPHPSPRGSPLTPKG-----TPVHTP 452
 DB 449 ARGGSPSTKQTLPSRGRGGGAGEQPPPPANRSTPLPGPGSPSSSGTPLHSLHTP 508
 QY 453 KESPACTNPPTPSSP--SVGGVPMRRLNSINKSFLGSPFRHRRKLQVPTPEENSLTP 510
 DB 509 RASPTGTPTTTPPSFGGCVGGAWRSLNSIRNSFLGSPFRHRRKLQVPTPEENSLTP 568
 QY 511 ESSPELAKSWFGNFISLEKEQIFVVLKDKPLSSIKADIVHAFSLPSLSHVSISQTSF 570
 DB 569 ESSPELAKSWFGNFISLKEEQIFVLKDKPLSSIKADIVHAFSLPSLSHVSISQTSF 628
 QY 571 RAEYKATGGPAVKFQVDFVITTEGEAQAQKE-----NGIYSVFTLLSGPSRRFRK 624
 DB 629 RAEYKASGGSPVQKPVRFQVDISSEGEPEPRDRDGGGGGIYSVFTLLSGPSRRFRK 688
 QY 625 VVETIOALLSTHDPPAAQHLSDDTN 650
 DB 689 VVETIOALLSTHDQPSVQALADEKN 714

RESULT 13

ABB62061

ID ABB62061 standard; Protein; 851 AA.

XX ABB62061;

XX 26-MAR-2002 (first entry)

XX Drosophila melanogaster polypeptide SEQ ID NO 12975.

XX Drosophila; developmental biology; cell signalling; insecticide;
 XX pharmaceutical.

OS Drosophila melanogaster.

XX WO200171042-A2.

XX PD 27-SEP-2001.

XX 23-MAR-2001; 2001WO-US09231.

XX 23-MAR-2000; 2000US-191637P.

XX 11-JUL-2000; 2000US-0614150.

XX (PEKE) PE CORP NY.

XX Venter JC, Adams M, Li PWD, Myers EW;

XX WPI; 2001-656860/75.

XX N-PSDB; ABL06164.

XX New isolated nucleic acid detection reagent for detecting 1000 or more
PT genes from Drosophila and for elucidating cell signalling and cell-cell
PT interactions -

XX Disclosure; SEQ ID NO 12975; 21pp + Sequence Listing; English.

XX The invention relates to an isolated nucleic acid detection reagent
CC capable of detecting 1000 or more genes from Drosophila. The invention is
CC useful in developmental biology and in elucidating cell signalling and
CC cell-cell interactions in higher eukaryotes for the development of
CC insecticides, therapeutics and pharmaceutical drugs. The invention
CC discloses genomic DNA sequences (AB16176-ABL30511), expressed DNA
CC sequences (ABL01840-ABL16175) and the encoded proteins
CC (ABB57737-ABB72072).

XX The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.

XX SQ Sequence 851 AA;

Query Match 56.4%; Score 1974; DB 22; Length 851;

Best Local Similarity 54.7%; Pred. No. 5.9e-137;

Matches 424; Conservative 85; Mismatches 134; Indels 132; Gaps 20;

QY 10 AQAQYGVYRLEKTLGKQGTGLVKLGVCVTCQKVAIKIVNREKLSVLMKVEREIAI 69

DB 9 AENCQFVGPYRLEKTLGKQGTGLVKLGVCVTCQKVAIKIVNREKLSVLMKVEREIAI 68

QY 70 LKLEHPHVLKLDVYENKLYLVLEHVSGGELFDYLVKKGRITPKRKFFQIIISAL 129

DB 69 MKLIDHPHVLGLSDVYENKLYLVLEHVSGGELFDYLVKKGRITPKRKFFQIIISAL 128

QY 130 DFCHSHSICHRLDKPENLLDKNNIRIADFGMASLOVGDLSLETSCSPHYACPEVIRG 189

DB 129 DFCHSHSICHRLDKPENLLDKNNIKIADFGMASLOVGDLSLETSCSPHYACPEVIRG 188

QY 190 EKYDGRKADVNSCGVILFALLVGLPFDLNLRLLEKVKRGVFMHPHIFPPDCQSLRG 249

DB 189 EKYDGRKADVNSCGVILFALLVGLPFDLNLRLLEKVKRGVFMHPHIFPPDCQSLRG 248

QY 250 MIEVDAARLLTLEHIOKHWY-IGGKNEPEPEQIPKRVQIRSLPSLEDIDPDVLDMSHS 308

DB 249 MIEVNDPRLTLAENRHPWVTAGGGELELELPMMEVYVTHVPTATAVDPDVNAICS 308

QY 309 LGCFRDRNKLQDLLSEBENCKMIFLLDRKERYPSQDEED--LPPRNE---IDPPR 362

DB 309 LGCFEKEKLIQELLSSHNTEKVIYFLLERKRRPALEDDDDIAKRSRSELDVDPFR 368

QY 363 KRVDSPMLN-----RHGKRPERKSEMEVLVSVD 390

DB 369 KRLDTCRINGTNAPYGOISEGSLTPRQAFNFRSYSTRNHRQRPSTTVTSVRS--SS 427

QY 391 GGSPVARRALEMAQ-----HGQSRISIGASGLSTPSLSS--- 427

DB 428 YHSFTRCNPSMSSAQQAANAIKSRPSPAAAGTRHSTYGRDR--SCHHSSVSRTSPSHSQK 485

QY 428 -----PRVTPHSPR-----GSP-----LTPPKGTP-----VHTPX 453

DB 486 SIEGDDVVVREPRIERRDSLQERGGSGPRDGCIPPGSPGSGNSGTSASPSVHHRA 545

QY 454 ES-----PAGTP-----NPTPPSSP--SVGGVPMRRLNSIKNSFLGSPFRHRRKL 497

DB 546 NSGPTTAAIIVNPNNGSPMNNSSPGMFGSPCNTPGGQWKTRLTNKNKNSFLGSPFRHRRKM 605

QY 498 QVPTPEMSNLTPESSPELAKSWFGNFISLEKEQIFVVKPKPLSSIKADIVHAFLSI 557

DB 606 QVSADE--VHLTPESSPELTKRSGFNLITTEDEFTTILVKGKIATVKAHLIHAFLSM 663

QY 558 PSLSHVISOTSFAEYKATG-GPAVFQKPVKQVDIT--YTEGGEAKENGIVSVTFTL 614

DB 664 AELSHSVSPTSFRVEYKNGNGFVNFQRHVKQVDLSAICKQGDADM---FALFTFTL 720

QY 615 LSGSPRRFRKRVETIOAQLLSTH-----DPP-----AAQHLSDTTNCMEMMTGRLS 660

DB 721 LSGNIRFRFRICEHIQSVCKRFGPGSPPTVTSVQAVSESSSCGVSSERLS 775

RESULT 14

ABBI1052

ID ABB11052 standard; peptide; 301 AA.

XX AC ABB11052;

XX DT 11-JAN-2002 (first entry)

XX DE Human HrPOPK-1 homologue, SEQ ID NO:1422.

XX Human; cytokine; cell proliferation; cell differentiation; growth factor;
KW haematopoiesis regulation; tissue growth; immunomodulator; activin;
KW inhibin; chemotaxis; chemokinesis; thrombolytic; oncogenesis;
KW proliferation; metastasis; cancer; tumour; haematopoietic disorder;
KW myeloid cell disorder; lymphoid cell disorder; asthma; arthritis;
KW chronic inflammatory condition; proliferative retinopathy;
KW atherosclerosis; coronary heart disease; arterial ischaemia;
KW bone disorder; osteoporosis; vascular growth disorder;
KW tissue regeneration; wound healing; infection; immune disorder;
KW cell culture; drug screening; gene therapy; antiinflammatory;
KW antiasthmatic; antiarthritic; haemostatic; antiarteriosclerotic;
KW cytostatic; osteopathic; vasotropic; cardiant; virucide; antibacterial;
KW antifungal; vulnery; antiulcer.

XX Homo sapiens.

OS WO200157188-A2.

XX PN 09-AUG-2001.

XX PD 05-FEB-2001; 2001WO-US03800.

XX 03-FEB-2000; 2000US-0496914.

XX 27-APR-2000; 2000US-0560875.

XX (HYSE-) HYSEQ INC.

XX Tang YT, Liu C, Drmanac RT;

XX WPI; 2001-457740/49.

XX N-PSDB; ABA08296.

XX Human proteins and DNA encoding sequences useful for preventing,
PT treating or ameliorating a medical condition in a mammalian subject
PT e.g. arthritis and cancer -

XX Claim 20; Page 140; 1963pp; English.

XX Sequences ABB10981-ABB12330 represent 1350 novel human polypeptides, and
CC sequences ABA08225-ABA09574 represent nucleic acids encoding them. The
CC invention also relates to vectors and recombinant host cells comprising a

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 26, 2003, 11:51:21 ; Search time 35.5987 Seconds
(without alignments)
4842.281 Million cell updates/sec

Title: US-10-054-579-2

Perfect score: 3497

Sequence: 1 MTSTGDKGGAHQAYGVGYPY.....TNCMEMTGRLSKCGIIPKS 668

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 830525 seqs, 258052604 residues

Total number of hits satisfying chosen parameters: 830525

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

SPTREMBL_23:**

- 1: sp_archaea:**
- 2: sp_bacteria:**
- 3: sp_fungi:**
- 4: sp_human:**
- 5: sp_invertebrate:**
- 6: sp_mammal:**
- 7: sp_mhc:**
- 8: sp_organelle:**
- 9: sp_phase:**
- 10: sp_plant:**
- 11: sp_rodent:**
- 12: sp_virus:**
- 13: sp_vertebrate:**
- 14: sp_unclassified:**
- 15: sp_rvirus:**
- 16: sp_bacteriap:**
- 17: sp_archaeap:**

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	3491	99.8	668	Q81WQ3	Q81WQ3 homo sapien
2	3015	86.2	503	O60843	O60843 homo sapien
3	2799	80.0	778	Q8TDC2	Q8TDC2 homo sapien
4	2785.5	79.7	794	Q8TDC3	Q8TDC3 homo sapien
5	2736.5	78.3	768	Q8NDD0	Q8NDD0 homo sapien
6	2599.5	74.3	715	Q8GJL4	Q8GJL4 homo sapien
7	1812.5	51.8	698	O61298	O61298 halocynthia
8	1743	49.8	914	Q19469	Q19469 caenorhabdi
9	1712	49.0	793	Q9VUV4	Q9VUV4 drosophila
10	1439.5	41.2	473	Q96AV4	Q96AV4 homo sapien
11	1248	35.7	701	Q95T82	Q95T82 drosophila
12	870.5	24.9	833	Q8SXX5	Q8SXX5 dictyosteli
13	796.5	22.8	745	Q15524	Q15524 homo sapien
14	796.5	22.8	755	Q96HB3	Q96HB3 homo sapien
15	794.5	22.7	691	Q96RG0	Q96RG0 homo sapien
16	792.5	22.7	722	11	O08679 rattus norv

17	791.5	22.6	888	11	Q8BR95	Q8BR95 mus musculu
18	790	22.6	401	8	Q98SC8	Q98SC8 guillardia
19	785.5	22.5	535	10	Q8RWD2	Q8RWD2 arabidopsis
20	784.5	22.4	514	10	Q9XF25	Q9XF25 glycine max
21	784.5	22.4	785	13	Q8QGV3	Q8QGV3 xenopus lae
22	784	22.4	505	10	Q92RJ1	Q92RJ1 oryza sativ
23	782.5	22.4	752	4	Q8NG37	Q8NG37 homo sapien
24	781	22.3	729	11	Q9JKE4	Q9JKE4 mus musculu
25	781	22.3	744	11	Q9JKE5	Q9JKE5 mus musculu
26	777.5	22.2	755	4	Q9P0L2	Q9P0L2 homo sapien
27	772.5	22.1	504	10	P931L3	P931L3 cucumis sat
28	771.5	22.1	752	11	Q8CIP4	Q8CIP4 mus musculu
29	771.5	22.1	926	4	Q9H0K1	Q9H0K1 homo sapien
30	771	22.0	512	10	P92968	P92968 arabidopsis
31	771	22.0	512	10	P92958	P92958 arabidopsis
32	768	22.0	797	11	Q8VHF0	Q8VHF0 rattus norv
33	766	21.9	626	5	Q95Q04	Q95Q04 caenorhabdi
34	764.5	21.9	793	11	O08678	O08678 rattus norv
35	764.5	21.9	795	11	Q8VHJ5	Q8VHJ5 mus musculu
36	764	21.8	931	11	Q8CFH6	Q8CFH6 mus musculu
37	763.5	21.8	511	10	Q40544	Q40544 nicotiana t
38	763	21.8	624	5	Q22068	Q22068 caenorhabdi
39	762	21.8	671	3	Q96W17	Q96W17 trichoderma
40	760.5	21.7	514	10	Q9M726	Q9M726 lycopersico
41	759.5	21.7	1371	4	Q9Y2K2	Q9Y2K2 homo sapien
42	754.5	21.6	942	5	Q8VMW9	Q8VMW9 haemonchus
43	754.5	21.6	1066	5	Q8VMX0	Q8VMX0 haemonchus
44	754	21.6	834	5	Q8VMX1	Q8VMX1 haemonchus
45	752.5	21.5	346	4	Q8NCV7	Q8NCV7 homo sapien

ALIGNMENTS

RESULT 1

Q81WQ3 ID Q81WQ3 PRELIMINARY; PRT; 668 AA.

AC Q81WQ3; 01-MAR-2003 (TREMREL. 23, Created)

DT 01-MAR-2003 (TREMREL. 23, Last sequence update)

DT 01-MAR-2003 (TREMREL. 23, Last annotation update)

DE Serine/threonine protein kinase isoform.

OS Homo sapiens (Human)

OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

OX NCBI_TaxID=9606;

RN [1]

RP SEQUENCE FROM N.A.

RC TISSUE=Brain tumor;

RA Guo J.H., Yu L.

RL Submitted (OCT-2002) to the EMBL/GenBank/DBDJ databases.

DR EMBL; AY166857; AAN87839.1; -

KW Kinase.

SQ SEQUENCE 668 AA; 74714 MW; 6AF8CB84FC48C07 CRC64;

Query Match	99.8%;	Score 3491;	DB 4;	Length 668;
Best Local Similarity	99.8%;	Pred. No. 1.9e-248;		
Matches 667;	Conservative 0;	Mismatches 1;	Indels 0;	Gaps 0;
Qy	1	MTSTGDKGGAHQAYGVGYPYRLEKTLGKGTGLVKLVGHVCTQKVAIKIVNREKLSSEVL	60	
Db	1	MTSTGDKGGAHQAYGVGYPYRLEKTLGKGTGLVKLVGHVCTQKVAIKIVNREKLSSEVL	60	
Qy	61	MKVERETAILKLTLEHPHVLKLDVYENKYLVLVLEHVSGLFDPYLVKGRLTPEARK	120	
Db	61	MKVERETAILKLTLEHPHVLKLDVYENKYLVLVLEHVSGLFDPYLVKGRLTPEARK	120	
Qy	121	FFRQIISALDFCHSHSICHRLDPENLLDEKKNIRIADFGMASLQVDSLETSCSPH	180	
Db	121	FFRQIISALDFCHSHSICHRLDPENLLDEKKNIRIADFGMASLQVDSLETSCSPH	180	
Qy	181	YACEVIRGEKYDGRKADVMSQGVILFALLVGLPFDNRLQLLEKVKRGVFMHPFIP	240	
Db	181	YACEVIRGEKYDGRKADVMSQGVILFALLVGLPFDNRLQLLEKVKRGVFMHPFIP	240	

```
Db 181 YACPEVIRGEKYDGRKADVMSGCVILFALLVGLALPFDNLLRQLLEKVKRGVFMHPFIP 240
Qy 241 PDCOSILRGHIEVDAAARLTLEHQHIWYIGGKNEPEPEQIPRKVQIRSLPSLEIDIP 300
Db 241 PDCOSILRGHSEVDAARLTLEHQHIWYIGGKNEPEPEQIPRKVQIRSLPSLEIDIP 300
Qy 301 DVLDSMHSGLGCFDRNKLQDLSEENQEKMIYFLLDRKERYPSQDEDLPRNEIDP 360
Db 301 DVLDSMHSGLGCFDRNKLQDLSEENQEKMIYFLLDRKERYPSQDEDLPRNEIDP 360
Qy 361 PRKRVDSPLNRHGRKRPERSMEVLVSDTGGSPVPARRAIEAQAQGRSISGASSGL 420
Db 361 PRKRVDSPLNRHGRKRPERSMEVLVSDTGGSPVPARRAIEAQAQGRSISGASSGL 420
Qy 421 STSPLSSPRVTPHPSRGSLPTPKGTVPHTPKESPAQTNPPTSPSPVGGVPMARLN 480
Db 421 STSPLSSPRVTPHPSRGSLPTPKGTVPHTPKESPAQTNPPTSPSPVGGVPMARLN 480
Qy 481 SIKNSFLGSRFRHRRKLQVPTPEMSNLTPSSPELAKKSWFGNFISLEKEEQIFVVIKD 540
Db 481 SIKNSFLGSRFRHRRKLQVPTPEMSNLTPSSPELAKKSWFGNFISLEKEEQIFVVIKD 540
Qy 541 KPLSSIKADIHAFSLSPSLSHSVISQTSFRAEYKATGGPAVFOKPVKFOVDITYTEGGE 600
Db 541 KPLSSIKADIHAFSLSPSLSHSVISQTSFRAEYKATGGPAVFOKPVKFOVDITYTEGGE 600
Qy 601 AQKENGIVSYVTFTLLSGPSRRFRKRVETIQALLSTHDPAAQHLSDTTNCMEWMTGRLS 660
Db 601 AQKENGIVSYVTFTLLSGPSRRFRKRVETIQALLSTHDPAAQHLSDTTNCMEWMTGRLS 660
Qy 661 KCGIIPKS 668
Db 661 KCGIIPKS 668

RESULT 2
O60843 PRELIMINARY; PRT; 603 AA.
ID O60843
AC O60843;
DT 01-AUG-1998 (TREMBlrel. 07, Created)
DT 01-AUG-1998 (TREMBlrel. 07, Last sequence update)
DT 01-MAR-2003 (TREMBlrel. 23, Last annotation update)
DE Putative serine/threonine protein kinase (Fragment).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX TISSUE=Brain;
RX MEDLINE=21064499; PubMed=11124703;
RA Stanchi F., Bertocco E., Toppo S., Dioguardi R., Simonati B.,
RA Cannata N., Zimbello R., Lanfranchi G., Valle G.;
RT "Characterization of 16 novel human genes showing high similarity to
RT yeast sequences."
RL Yeast 18:69-80(2001).
CC -1- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
DR EMBL; AJ006701; CAA07196.1; -.
DR HSSP; Q63450; 1A06.
DR Genew; HGNC:11405; STK29.
DR InterPro; IPR000719; Prot_kinase.
DR InterPro; IPR002290; Ser_thr_pkinase.
DR Pfam; PF00069; pkinase; 1.
DR PRODOM; PD000001; Prot_kinase; 1.
DR PROSITE; PS00011; PROTEIN_KINASE_DOM; 1.
DR PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
KW ATP-binding; Kinase; Serine/threonine-protein kinase; Transferase.
FT NON TER
SQ SEQUENCE 603 AA; 67401 MW; B02C5D678F8E96 CRC64;
```

Query Match 86.2%; Score 3015; DB 4; Length 603;
Best Local Similarity 99.7%; Pred. No. 1.8e-213;
Matches 574; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

```
Qy 72 LLEHPIVLKLDHVENKKYLVLVLEHVSGLGELFDYLKKGRLLTPKEARKFFRQIISALDF 131
Db 1 LLEHPIVLKLDHVENKKYLVLVLEHVSGLGELFDYLKKGRLLTPKEARKFFRQIISALDF 60
Qy 132 CHSHSICHRLDKPENLLDDEKNINRIADFGMASLQVGDLSLETSCGSHYACPEVIRGEK 191
Db 61 CHSHSICHRLDKPENLLDDEKNINRIADFGMASLQVGDLSLETSCGSHYACPEVIRGEK 120
Qy 192 YDGRKADVMSGCVILFALLVGLALPFDNLLRQLLEKVKRGVFMHPFIPDQCOSILRGH 251
Db 121 YDGRKADVMSGCVILFALLVGLALPFDNLLRQLLEKVKRGVFMHPFIPDQCOSILRGH 180
Qy 252 EYDAARRLTLEHQHIWYIGGKNEPEPEQIPRKVQIRSLPSLEIDIPDVLDSMHSGLG 311
Db 181 EYDAARRLTLEHQHIWYIGGKNEPEPEQIPRKVQIRSLPSLEIDIPDVLDSMHSGLG 240
Qy 312 FRDRNKLQDLSEENQEKMIYFLLDRKERYPSQDEDLPRNEIDPPRKRVDSPLN 371
Db 241 FRDRNKLQDLSEENQEKMIYFLLDRKERYPSQDEDLPRNEIDPPRKRVDSPLN 300
Qy 372 RHGRKRPERSMEVLVSDTGGSPVPARRAIEAQAQGRSISGASSGLSTPLSSPRVT 431
Db 301 RHGRKRPERSMEVLVSDTGGSPVPARRAIEAQAQGRSISGASSGLSTPLSSPRVT 360
Qy 432 PHPSRGSLPTPKGTVPHTPKESPAQTNPPTSPSPVGGVPMARLNSIKNSFLGSPR 491
Db 361 PHPSRGSLPTPKGTVPHTPKESPAQTNPPTSPSPVGGVPMARLNSIKNSFLGSPR 420
Qy 492 FHRRKLQVTPPEMSNLTPSSPELAKKSWFGNFISLEKEEQIFVVIKDKPLSSIKADIV 551
Db 421 FHRRKLQVTPPEMSNLTPSSPELAKKSWFGNFISLEKEEQIFVVIKDKPLSSIKADIV 480
Qy 552 HAFSLSPSLSHSVISQTSFRAEYKATGGPAVFOKPVKFOVDITYTEGGEAQKENGIVSYVT 611
Db 481 HAFSLSPSLSHSVISQTSFRAEYKATGGPAVFOKPVKFOVDITYTEGGEAQKENGIVSYVT 540
Qy 612 FTLLSGPSRRFRKRVETIQALLSTHDPAAQHLSD 647
Db 541 FTLLSGPSRRFRKRVETIQALLSTHDPAAQHLSE 576

RESULT 3
Q8TDC2 PRELIMINARY; PRT; 778 AA.
ID Q8TDC2
AC Q8TDC2;
DT 01-JUN-2002 (TREMBlrel. 21, Created)
DT 01-JUN-2002 (TREMBlrel. 21, Last sequence update)
DT 01-MAR-2003 (TREMBlrel. 23, Last annotation update)
DE Protein kinase-like protein.
OS Homo sapiens (Human)
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA She X.Y., Yu L., Guo J.H.;
RL Submitted (FEB-2002) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
DR EMBL; AF479827; AAL87698.1; -.
DR HSSP; P24941; 1BUH.
DR InterPro; IPR000719; Prot_kinase.
DR InterPro; IPR002290; Ser_thr_pkinase.
DR InterPro; IPR000449; UBA_domain.
DR Pfam; PF00069; pkinase; 1.
DR PRODOM; PD000001; Prot_kinase; 1.
DR SMART; SM00220; S_TK; 1.
DR PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
DR PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
KW ATP-binding; Kinase; Serine/threonine-protein kinase; Transferase.
SQ SEQUENCE 778 AA; 85086 MW; 8D1818D4E54398BB CRC64;
```

Query Match	80.0%;	Score 2799;	DB 4;	Length 778;
Best Local Similarity	75.1%;	Pred. No. 2.1e-197;		
Matches 548;	Conservative 54;	Mismatches 48;	Indels 80;	Gaps 8;
QY	1	MTSTGDKDGA-----	QAAQYVGPVRLEKTLGKGOTGLVGLGVHCVTCQKV	45
DB	1	MSSAKEGGGGSPAYHLPHPHPHPQHAQYVGPVRLEKTLGKGOTGLVGLGVHCVTCQKV	60	
QY	46	AIKIVNREKLSYLMVKVEREIALKUIEHPHVLKJHDVYENKKYLVLVLEHVS GGELFD	105	
DB	61	AIKIVNREKLSYLMVKVEREIALKUIEHPHVLKJHDVYENKKYLVLVLEHVS GGELFD	120	
QY	106	YLVKKGLTPKEARKFPRQIISALDFCHSHSICHRLDKPENLILDEKKNIRIADFGNASL	165	
DB	121	YLVKKGLTPKEARKFPRQIVSALDFCHSYSICHRLDKPENLILDEKKNIRIADFGNASL	180	
QY	166	QVGDLSLETSCGSPHYACPEVIRGEIKYDGRKADVWSCGVIIFALLVGNALPDPDDNLQQLL	225	
DB	181	QVGDLSLETSCGSPHYACPEVIRGEIKYDGRADVMWSCGVIIFALLVGNALPDPDDNLQQLL	240	
QY	226	EKVKGVFHMPHFTPPQCQSLLRGMIEVDAAARLTLEHIQKHIWIYIGKNEPEP-EQPIIP	284	
DB	241	EKVKGVFHMPHFTPPQCQSLLRGMIEVEPEKRLSLEIQKHFWYLGCKHEPDPCLSEFAP	300	
QY	285	-RKVOISRLSLEDIDPDVLDMSHSLGCFRDNKLLQDLLSEESNOEMTYFILLDRKER	343	
DB	301	GRRVVAFLSPNGELDPVLESMSALGCFRERLHRELSEENQEMTYFILLDRKER	360	
QY	344	YPSQEDDLPPRNIDPPRKVDVSPMLNRHGKRRPERKSMELSVTD---CGSPVPARRA	400	
DB	361	YPSCEDQLPPRNVDPPRKEVDSPMLSRHGKRRPERKSMELSVITDAGGGSPVPTERRA	420	
QY	401	LEMAHQQRGRSISGASGLSTSPLSGR-----	429	
DB	421	LEMAHQQRGRSVSGASTGLSSSLSPSPSVFSPGPGAGDARGGGSPSTQTPLPS	480	
QY	430	-----VTPHPSPRGSPLPKPG-----	468	
DB	481	RGPRGGAGAQPPPPPSARSTPLPGPGSPRSSGGTPLHSPHTPRASGTPTGTPPTPPSP	540	
QY	469	--SYGVGPVRRALNSIKNSFLGSPRHRKLVQPTPEMSNLTPESSPELAKSWFNGFI	526	
DB	541	GGGVGGAARSLNSIRNSFLGSPRHRKLVQPTPEMSNLTPESSPELAKSWFNGFI	600	
QY	527	SLEKEQIFVVIKDKPLSSIKADIVHAFLSIPSLSHSVISQTSFRAZYKATGGPAVFOKP	586	
DB	601	SLDKEQIFVLKDKPLSSIKADIVHAFLSIPSLSHSVISQTSFRAZYKATGGPAVFOKP	660	
QY	587	VKFOVDLITYTEGGBAQKE-----NGIYSVTFLLSGPSRRFRKRVVETIOAQLLSTHDP	640	
DB	661	VRQVDLISSEGGPEPSPRDGSGGGGIYSVTFLLSGPSRRFRKRVVETIOAQLLSTHDP	720	
QY	641	AAHLSDTN	650	
DB	721	SVOALADEKN	730	

RESULT 4	
Q8TDC3	
ID	PRELIMINARY;
AC	PRT; 794 AA.
Q8TDC3	
DT	01-JUN-2002 (TrEMBLrel. 21, Created)
DD	01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DE	01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE	Putative serine/threonine protein kinase.
OS	Homo sapiens (Human).
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC	Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
OX	NCBI_TaxID=9606;
RN	[1]
RP	SEQUENCE FROM N.A.
RP	TISSUE=Brain;
RC	She X.Y., Guo J.H., Yu L.;

Submitted (FEB-2002) to the EMBL/GenBank/DBJP databases.
-1- SIMILARITY: ANALOGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
EMBL: AF479826; AAU87697.1; --
HSP: P24941; 1BUH.
InterPro: IPR000719; prot_kinase.
InterPro: IPR002290; Ser_Thr_kinase.
InterPro: IPR000449; UBA_Domain.
Pfam: PF00069; pkinase; 1.
ProDom: PD000001; prot_kinase; 1.
SMART: SM00220; S_TKc_1.
PROSITE: PS00107; PROTEIN_KINASE_ATP; 1.
PROSITE: PS00011; PROTEIN_KINASE_DOM; 1.
PROSITE: PS00108; PROTEIN_KINASE_ST; 1.
ATP-binding; Kinase; Serine/threonine-protein kinase; Transferase.
Sequence 794 AA: 86753 MW: 5D0395B0E61AEF77 CFC64.
SEQUENCE

Query Match	79.7%;	Score	2785.5;	DB	4;	Length	794;
Best Local Similarity	76.9%;	Pred.	No. 2.1e-196;				
Matches	542;	Conservative	53;	Mismatches	45;	Indels	65;
							Gaps
Qy	11	QHAQYVGPVLEKTLGKGTGLVKLVGHCVTCOKVAIKIVNREKLSVSLMKVERETAIL	70	:	:	:	:
Db	42	RHAQYVGPVLEKTLGKGTGLVKLVGHCVTCOKVAIKIVNREKLSVSLMKVERETAIL	101	:	:	:	:
Qy	71	KLIEHPVLKLDHVDYENKYLVLVLEHVS GGELFDYLVKKGRLLTPPEAKRFFRQIIISALD	130	:	:	:	:
Db	102	KLIEHPVLKLDHVDYENKYLVLVLEHVS GGELFDYLVKKGRLLTPPEAKRFFRQIVSALD	161	:	:	:	:
Qy	131	FCSHSICHRLDLPENLLDDEKKNIRIADFGMASLQVGDLSLLETSCGSPHYACPEVIRGE	190	:	:	:	:
Db	162	FCHSYGICHRLDLPENLLDDEKKNIRIADFGMASLQVGDLSLLETSCGSPHYACPEVIKGE	221	:	:	:	:
Qy	191	KYDGRKADVMSCGVILFALLVGLPFPDDNNLRQLLEKVKRGVFMHPHFIIPDQCOSLLRGM	250	:	:	:	:
Db	222	KYDGRADVMSCGVIIFALLVGLPFPDDNNLRQLLEKVKRGVFMHPHFIIPDQCOSLLRGM	281	:	:	:	:
Qy	251	IEVDAARRLTLEHIOKHVIWYIGGKNPEP-EQPIP-RKVOIRSLPSLEDIDPDVLDMSHS	308	:	:	:	:
Db	282	IEVEPEKRLSLEQIQKHPWYLGCKHEPDPCLPEPAPGRRVAMRSLNPSNGELDPDVLESMA	341	:	:	:	:
Qy	309	LGCFDRNKLLOLLSEENOEKMIYFLLLDRKERYPSQEDLPPEINETDPPRKRYDVP	368	:	:	:	:
Db	342	LGCFDRERLHRELSEENOEKMIYLLLLDRKERYPSCEDQLPPANDVDPFRKRYDVP	401	:	:	:	:
Qy	369	MLNRHGKRPERKSMEVLSVTD---CGSPVPARRAIEMAHQGRSRISGASSGLSTSP	425	:	:	:	:
Db	402	MLSRHGKRPERKSMEVLSITDAGGGSPVPTRRALEMAHQSRKRSVSGASTGLSSSP	461	:	:	:	:
Qy	426	SSPR-----	444	:	:	:	:
Db	462	SSPRSPVSPSPGAGDBARGGSPSTKQTULPSRGPGGGAGEOPPPSPASRSTPLPGP	521	:	:	:	:
Qy	445	KG-----TPVHTPKESPAGTNPNTPPSSP--SVGGVPWRARLNSIKNSFLGSPR	491	:	:	:	:
Db	522	PGSPRSSGGTPLHSLPLHTPRASPTGTPGTTPPSPGGVGGAAWRSLNSIRNSFLGSPR	581	:	:	:	:
Qy	492	FHRKLOVPTPEMGNLTPESSPELAKKSWFGNFI SLEKEEQI FVVIKDKPLSSIKADIV	551	:	:	:	:
Db	582	FHRKKNQVPTAEEMSSLTPESSPELAKRSWFGNFI SLDKEEQIFLVLKDKPLSSIKADIV	641	:	:	:	:
Qy	552	HAFSLSPSLSHSVISOTSFRAEYKATGGPAVFOKPVKFQVDITYTGEGBAQKE-----N	605	:	:	:	:
Db	642	HAFSLSPSLSHSVLSOTSFRAEYKASGGSPVOKPVRFQVDI SSSGGPEPSPRDGGSGG	701	:	:	:	:
Qy	606	GIYSVTFTLTLLSGSPRRFKRVVETIQALLSTHDPAPAAQHLSDTTN	650	:	:	:	:
Db	702	GIYSVTFTLTIGSPRRFKRVVETIQALLSTHDPQSPVQALADEKN	746	:	:	:	:

RESULT 5

Q8ND00

ID Q8ND00

AC Q8ND00;

PRELIMINARY;

PRT; 768 AA.

DT 01-OCT-2002 (TrEMBLrel. 22, Created)
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Hypothetical protein (fragment).
GN DXFZP54781613.
OS Homo sapiens (Human).
OC Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Brain;
RA Koehrer K., Beyer A., Mewes H.W., Weil B., Wiemann S.;
RL Submitted (JUL-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AL834275; CAB38950.1; -
DR InterPro; IPR000719; Prot_kinase.
DR InterPro; IPR002290; Ser_thr_kinase.
DR InterPro; IPR001245; Tyr_kinase.
DR InterPro; IPR000449; UBA domain.
DR Pfam; PF00069; pkinase; 1.
DR ProDom; PD000001; Prot_kinase; 1.
DR SMART; SM00220; S_TKC; 1.
DR SMART; SM00219; Tyrc; 1.
DR PROSITE; PS00107; PROTEIN KINASE ATP; 1.
DR PROSITE; PS00111; PROTEIN KINASE DOM; 1.
DR PROSITE; PS00108; PROTEIN KINASE ST; 1.
KW Hypothetical protein; ATP-binding; Transferase.
FT NON_TER 1
SQ SEQUENCE 768 AA; 83962 MW; 1P954ADF62B89A CRC64;
Query Match 76.3%; Score 2736.5; DB 4; Length 768;
Best Local Similarity 76.7%; Pred. No. 8e-193;
Matches 534; Conservative 52; Mismatches 45; Indels 65; Gaps 7;
QY 20 RLEKTLKGGTGLVGLVGHVCTQKVAIKIVNREKLSVLMKVVERIAIKLIEHPHVL 79
DB 1 RLEKTLKGGTGLVGLVGHVCTQKVAIKIVNREKLSVLMKVVERIAIKLIEHPHVL 60
QY 80 KLHDVYENKYLVLVLEHVSGLGFDLVKKGRLTPKEARKFQRIISALDFCHSHSICH 139
DB 61 KLHDVYENKYLVLVLEHVSGLGFDLVKKGRLTPKEARKFQRIISALDFCHSHSICH 120
QY 140 RDLKPNLLDEKNNIRIADFGMASLQVDSLLTSCGSPHYACPEVIRGEKDGKADV 199
DB 121 RDLKPNLLDEKNNIRIADFGMASLQVDSLLTSCGSPHYACPEVIRGEKDGKADV 180
QY 200 WSCGVILFALLVGLPFDNDNLRLLEKVKRGVFMHPHFIPDCQSLLRGMIEVDAAARL 259
DB 181 WSCGVILFALLVGLPFDNDNLRLLEKVKRGVFMHPHFIPDCQSLLRGMIEVDAAARL 240
QY 260 TLEHIQKHWYIGKNEPEP-EQIP-RKVQIRSLPSLEDIPDVLDSMHSGLGCFDRNK 317
DB 241 SLEQIQKHPWYLGKHPDPCLEPAPARRVAMRSLPSNGELDPDVLDSMHSGLGCFDRNK 300
QY 318 LLODLLSEENQEKMIYFLLLDRKERYPSQDEDLPPRNEIDPPRKRVDSPMLNRHGKR 377
DB 301 LHRELSEENQEKMIYFLLLDRKERYPSQDEDLPPRNDVDPKRKVDSPMLNRHGKR 360
QY 378 PERKSMELSVTD---GGSPVAPARALTEMAHQGRSISGASGLSTPSLSSPR----- 429
DB 361 PERKSMELSVTDAGGGGSPVTPTRALEMAHQGRSISGASGLSTPSLSSPRSVFS 420
QY 430 -----VTPHSPRSGSLPTPKG----- 446
DB 421 FSPFPGADGARGGSPSTKQTQLPSRCPGPGGAGEQPPPSARSTPLPGPSPRSGG 480
QY 447 ----TPVHTPKESPAGTNPPTPSSP--SVGVWPVAPRLNISKNSFLGSPRFRKRLQVP 500
DB 481 TPLHSLPTPRASPTGPTGTPPPSPGGVGAARSLNIRSNFLGSPRFRKRLQVP 540
QY 501 TPBMSNLTPSSPELAKSHFGNFIISLEKEQFVVIKDKPLSIKADIIVHAFSLPSL 560
DB 541 TAEMSSLTSPSSPELAKRSWFGNFIISLDEKEQIFVLKDKPLSLSIKADIIVHAFSLPSL 600

QY 561 SHSVISOTSFRAEYKATGPAVQKPAVQKFOVDITYTEGEAKKE-----NGIYSVTFTL 614
DB 601 SHSVISQTSFRAEYKASGPSVQKPAVQKFOVDISSESGEPSPRRDGGSGGGGIYSVTFTL 660
QY 615 LSGPSRRRKRVVETITQALLSTHDPAPAAQHLSDTTN 650
DB 661 ISGPSRRRKRVVETITQALLSTHDPAPAAQHLSDTTN 696
RESULT 6
Q96JL4 PRELIMINARY; PRT; 715 AA.
AC Q96JL4;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Hypothetical protein KIAA1811 (fragment).
GN KIAA1811.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Brain;
RA Nagase T., Nakayama M., Nakajima D., Kikuno R., Ohara O.;
RL MEDLINE=21245130; PubMed=11347906;
RT "Prediction of the coding sequences of unidentified human genes. XX.
RT The complete sequences of 100 new cDNA clones from brain which code
RT for large proteins in vitro."
RL DNA Res. 8:85-95(2001).
CC -!- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
DR EMBL; A5058714; BAB47440.1; -
DR HSSP; P24941; IBUH.
DR InterPro; IPR000719; Prot_kinase.
DR InterPro; IPR002290; Ser_thr_kinase.
DR InterPro; IPR000449; UBA domain.
DR Pfam; PF00069; pkinase; 1.
DR ProDom; PD000001; Prot_kinase; 1.
DR SMART; SM00220; S_TKC; 1.
DR PROSITE; PS00111; PROTEIN KINASE DOM; 1.
DR PROSITE; PS00108; PROTEIN KINASE ST; 1.
KW Hypothetical protein; ATP-binding; Kinase;
KW Serine/threonine-protein kinase; Transferase.
FT NON_TER 1
SQ SEQUENCE 715 AA; 78499 MW; B90F6EE115C418A5 CRC64;
Query Match 74.3%; Score 2599.5; DB 4; Length 715;
Best Local Similarity 76.0%; Pred. No. 8.9e-183;
Matches 507; Conservative 51; Mismatches 44; Indels 65; Gaps 7;
QY 49 IVNREKLSVLMKVVERIAIKLIEHPHVLKLDVYENKYLVLVLEHVSGLGFDLV 108
DB 1 IVNREKLSVLMKVVERIAIKLIEHPHVLKLDVYENKYLVLVLEHVSGLGFDLV 60
QY 109 KKGRLTPKEARKFQRIISALDFCHSHSICHRLDKPENLLDEKNNIRIADFGMASLQV 168
DB 61 KKGRLTPKEARKFQRIISALDFCHSHSICHRLDKPENLLDEKNNIRIADFGMASLQV 120
QY 169 DSLLTSCGSPHYACPEVIRGEKDGKADVMSGCVILFALLVGLPFDNDNLRLLEKV 228
DB 121 DSLLTSCGSPHYACPEVIRGEKDGKADVMSGCVILFALLVGLPFDNDNLRLLEKV 180
QY 229 KRGVFMHPHFIPDCQSLLRGMIEVDAAARLTLEHIQKHWYIGKNEPEP-EQIP-RK 286
DB 181 KRGVFMHPHFIPDCQSLLRGMIEVDAAARLTLEHIQKHWYIGKNEPEP-EQIP-RK 240
QY 287 VQIRSLPSLEDIPDVLDSMHSGLGCFDRNKRLQDLLSEENQEKMIYFLLLDRKERYPS 346
DB 241 VAMRSLPSNGELDPDVLDSMHSGLGCFDRNRHLRELSEENQEKMIYFLLLDRKERYPS 300
QY 347 QEDELPPRNEIDPPRKRVDSPMLNRHGKRPERKSMELSVTD---GGSPVAPARAIEM 403


```
Db 301 CEDQDLPPNDVDPPKRVDSFVMSLRSRGKRRPERKSMELVLSITDAGGGSPVPTRRALM 360
Qy 404 AOHQORSISGASSGLSTPLSPR----- 429
Db 361 AOHQORSISGASTGLSSLSLSPRSPVFSPEPGAGDEARGGGSPSTKOTLPSRGP 420
Qy 430 -----VTPHPSRPGSLPTPKG-----TPVHTPKESPAGTTPNTPPSSP--S 469
Db 421 RGGGAGQPPPSARSPTLPFGPPSPRSGTGTPHSLHTPRASPTGTPTTTPPSPGGG 480
Qy 470 VGVFWARLINSIKNSFLGSPFRHRRKQVPTPEMSNLTPESSPELAKSWFGNFTSLE 529
Db 481 VGGAAWRSRLNSIRNSFLGSPFRHRRKQVPTAEEMSLTPESSPELAKSWFGNFTSLD 540
Qy 530 KEEQIFVVKDKPLSSIKADIVHAFSLTPSLSHSVISQTSFRAEYKATGGPAVFKQKVP 589
Db 541 KEEQIFVLKDKPLSSIKADIVHAFSLTPSLSHSVISQTSFRAEYKATGGPVSFQKVP 600
Qy 590 QVDITYTEGGEAQKE-----NGIYSVTFTLLSGSPRRFRKRVETIOALLSTHDPAAQ 643
Db 601 QVDISSSEGGPSPRRDGGGGIYSVTFTLLSGSPRRFRKRVETIOALLSTHDPQSVQ 660
Qy 644 HLDSTTN 650
Db 661 ALADEKN 667

RESULT 7
O61298 PRELIMINARY; PRT; 698 AA.
ID AC O61298;
AC O61298;
DT 01-AUG-1998 (TrEMBLrel. 07, Created)
DT 01-AUG-1998 (TrEMBLrel. 07, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE HRPPOK-1 protein.
GN HRPPOK-1.
OS Halocynthia roretzi (Sea squirt).
OC Eukaryota; Metazoa; Chordata; Urochordata; Ascidiacea;
OC Stolidobranchia; Pyrosidae; Halocynthia.
OX NCBI_TaxID=7729;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Egg;
RX MEDLINE=98440280; PubMed=9767157;
RA Sasakura Y., Ogasawara M., Makabe K.W.;
RT "Maternally localized RNA encoding a serine/threonine protein kinase
in the ascidian, Halocynthia roretzi.";
RL Mech. Dev. 76:161-163(1998).
CC -1- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
DR EMBL; AB014885; BAA28663.1; -.
DR HSSP; O63450; 1A06
DR InterPro; IPR000719; Prot_kinase.
DR InterPro; IPR002290; Ser_thr_kinase.
DR InterPro; IPR001245; Tyr_kinase.
DR Pfam; PF00069; pkinase; 1.
DR PRINTS; PR00109; TYRKINASE.
DR ProDom; PD000001; Prot_kinase; 1.
DR SMART; SM00220; S_TK; 1.
DR PROSITE; PS00107; PROTEIN KINASE ATP; 1.
DR PROSITE; PS00011; PROTEIN KINASE DOM; 1.
DR PROSITE; PS00108; PROTEIN KINASE ST; 1.
KW ATP-binding; Kinase; Serine/threonine-protein kinase; Transferase.
SQ SEQUENCE 698 AA; 78607 MW; B9B1751E83200BF CRC64;

Query Match 51.8%; Score 1812.5; DB 5; Length 698;
Best Local Similarity 55.7%; Pred.No. 7.2e-125;
Matches 371; Conservative 103; Mismatches 125; Indels 67; Gaps 16;

Qy 11 QHAGYVGYRLEKTLGKQTLVLKGLVCHVTCQVAKIVNREKLSVLMKVEREAIL 70
Db 6 QPGQYVGYRLEKTLGKQTLVLKGLVCHVTCQVAKIVNREKLSVLMKVEREAIL 65
```

```
Qy 71 KLIEHPVLKLDHVVYENKKYLVILVLEHVSGLFDYLVKKGLRTPKEARKFRQIISALD 130
Db 66 KLIEHPHILGLHDVYENKKYLVILVLEHVSGLFDYLVKKGLRTPKEARKFRQIISAVD 125
Qy 131 FCHSHSICHRLDKPENLLLDKKNIRIADFGMASLQVGDLSLETSCGSPHYACPEVIRGE 190
Db 126 YCHNHNVCHRLDKPENLLLDKKNIKVADFGMASLQVGEFLLSETSCGSPHYACPEVIRGE 185
Qy 191 KYDGRADVWSCGVILFALLVGNALPEDDNLRLQLLEKVKRGVFMHPFIPDPQCSLLRGM 250
Db 186 RYDGRADVWSCGVILFALLVGNALPEDDNLRLQLLEKVKRGVYHIFHVPVDDAQNLLRGM 245
Qy 251 IEYDAARRLTLEHIOKHIVIGGKNE-----PEPEQIPRKVQIRSLPSLEDIDPDVL 303
Db 246 IDVVPDKRLSLQOVQLGHPWMPGNSVSGVLTPTDPVPV-----IDCVPLPEESVDPDL 302
Qy 304 DSMHSLGCFRDRNKLQDLLESENEQEKMIYFLLDRKERYPS-QBEDDLPPRNEIDPPR 362
Db 303 ASMTSLGCFNCKEKLKLNLTTEQNTKVVVYMLLRKKRYPSPFDDADSLPCKHPDAPR 362
Qy 363 KRVD--SPMLNRHG-----KRPERK-SMEVLSTVDTGGSPVARRAIEAQAHCQRSSIS 414
Db 363 KRVDSTSLSSNGDDWCNPIPQRKMSAESLCLTSSSPFLSRK--KSTETHQRQSQT 420
Qy 415 GASSEL-----STSPLSRPTVPHSPRGSPLTPKGTPEVH--TPKESPACTPNPTTP 465
Db 421 GESNRLVCNISDTKAESKRI-----NGTPVRGTCSSNQVPQINTP 466
Qy 466 SPSVGVGVPWRAFLNSIKNSFLGSPFRHRRKQVPTPEMSNLTPESSPELAKSWFGNF 525
Db 467 ASDN----PWRQRLASLKNKTFMGSPFRHRRKQAPSSDEVEN-QGNSSSELSKRSWFGNF 521
Qy 526 ISLE-----KEEQIFVVKOKPLSSIADIVHAFSLTPSLSHSVISQTSFRAEYKATG 578
Db 522 MSSRYSSTEHCDLPYAIYKNTLNSVKSELVHAFSLTPSLSHSVISQTSFRAEYKATG 581
Qy 579 --GPAYF-QKPVKQFQVDITVTEGGEAQKEN-----GIYSVTFTLLSGSPRRFRKRVVE 627
Db 582 TSTTSVFHQRSIKFQVDIIQHSLSLDQENGKPKSSQTSVTSFTIAFSLGPIRYKRVLE 641
Qy 628 TIQAQL 633
Db 642 LLQMQM 647

RESULT 8
Q19469 PRELIMINARY; PRT; 914 AA.
ID AC Q19469; Q9BMN6;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-OCT-2001 (TrEMBLrel. 18, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE F15A2.6 protein (Serine/threonine kinase SAD-1).
GN F15A2.6.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditidae;
OC Rhabditidae; Peloderinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
RP SEQUENCE FROM N.A.
RA Gregory J.;
RL Submitted (MAR-1996) to the EMBL/GenBank/DBJ databases.
[2]
RN SEQUENCE FROM N.A.
RX MEDLINE=99069613; PubMed=9851916;
RA none;
RT "Genome sequence of the nematode C.elegans: A platform for
investigating biology.";
RL Science 282:2012-2018(1998).
[3]
RN SEQUENCE FROM N.A.
RA Crump J.G., Zhen M., Jin Y., Bargmann C.;
RT "The SAD-1 kinase regulates presynaptic vesicle clustering.";
```



```

RA Williams S.M., Zaveri J.S., Smith H.O., Venter J.C., Rubin G.M.;
RT "Sequencing of Drosophila melanogaster genome.";
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A..
RA Misra S., Crosby M.A., Matthews B.B., Bayraktaroglu L., Campbell K.,
RA Hradecky P., Huang Y., Kaminker J.S., Prochuk S.E., Smith C.D.,
RA Tupy J.L., Bergman C., Bernier J.W., Carlson J.W., Celisner S.E.,
RA Clump M., Drysdale R., Emmert D., Frise E., de Grey A., Harris N.,
RA Krommiller B., Marshall B., Millburn G., Richter J., Russo S.,
RA Searle S.M.J., Smith E., Shu S., Smutniak F., Whitfield B.,
RA Ashburner M., Gelbart W.M., Rubin G.M., Mungall C.J., Lewis S.E.;
RT "Annotation of Drosophila melanogaster genome.";
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
RN [4]
RP SEQUENCE FROM N.A..
RA Adams M.D., Celisner S.E., Gibbs R.A., Rubin G.M., Venter C.J.;
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
RN [5]
RP SEQUENCE FROM N.A..
RA FlyBase;
RL Submitted (SEP-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AE003529; RA049569.2; -.
DR HSP; Q63450; IAO6.
DR FlyBase; FBgn0036544; CG6114.
DR InterPro; IPR000719; Prot kinase.
DR InterPro; IPR002290; Ser Thr kinase.
DR InterPro; IPR001245; Tyr kinase.
DR Pfam; PF00069; pkinase; 1.
DR ProDom; PD000001; Prot_kinase; 1.
DR SMART; SM00220; S_TKc; 1.
DR SMART; SM00219; TyrK; 1.
DR PROSITE; PS00111; PROTEIN_KINASE_DOM; 1.
DR PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
KW ATE-binding; Transferase.
SQ SEQUENCE 793 AA; 87692 MW; DCOF371B489FEF48 CRC64;

Query Match 49.0%; Score 1712; DB 5; Length 793;
Best Local Similarity 51.0%; Pred. No. 2.1e-117;
Matches 370; Conservative 80; Mismatches 133; Indels 142; Gaps 19;

QY 70 LKLEHPLVHLKHDVYENKLYLVLEHVSQGLDFYLVKGRITPKPKARFFRQIISAL 129
Db :||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
QY 1 MKLIDHPVLGLSDVYENKLYLVLEHVSQGLDFYLVKGRITPKPKARFFRQIISAL 60
Db :||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:

QY 130 DFCSHSICHRLDKPENLLDEKNIRIADFGMASLQVDSLETSCGSHYACPEVIRG 189
Db :||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
61 DFCSHSICHRLDKPENLLDEKNIKIADFGMASLQVDSLETSCGSHYACPEVIRG 120
Db :||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:

QY 190 EKYDGRKADVMSCGVLFFALLVGLPFDNDNLRLQLLEKVRGVFHPHFIPDCQSLRG 249
Db :||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
121 EKYDGRKADVMSCGVLVYALLVGLPFDNDNLRLQLLEKVRGVFHPHFIPDCQSLRG 180
Db :||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:

QY 250 MIEVDAARLLEHTQKIHW-IGKNPEPEQPIPRKVQIRSLPSIEDIPDVLDSMHS 308
Db :||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
181 MIEVNPDRLLAEINRHPWVTAGKGLELELPMMEVVQTHVPTATAVDPDLNATCS 240
Db :||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:

QY 309 LGCFRDRNKLQDLLEENQEKMIYFLLDRKERYPSQED---LPFRNE---IDPPR 362
Db :||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
241 LGCFAKEKELQELSSSHNTEKVIYFLLERKRRPALEDDDEIAQKRSLEDAVPPR 300
Db :||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:

QY 363 KRVDSPMLN-----RHGKRRRPERKSMELSVTD 390
Db :||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
301 KRLDTCRINGTNAPSYGOISEGSLTPRQAFNFRSYSTRHQERSPTTYSVRS--SS 359
Db :||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:

QY 391 GGSVPARRAEMAQ-----HGQSRISGASGLSTSPSS--- 427
Db :||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
360 YHSPTRCNPSMSSAQQAANASRPSPAAGTRHSTYGDRLR--SGHSSSVRSTSHSQK 417
Db :||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
428 -----PRVTPHPSR-----GSP-----LPTPKGTP----- 448
Db :||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
418 SIEGDVVVRPREIRERDLSLQERGGSPDRGDCGIPGSPGNGSGSTASPSVHRA 477

```

```
Db 361 DISSEGEPEPRRDGSGGGIYVTFLLISGPSRRFKRVVETIQALLSTHDQPSVAL 420
Qy 646 SDTTN 650
Db 421 ADEKN 425

RESULT 11
Q95T82 PRELIMINARY; PRT; 701 AA.
AC Q95T82;
DT 01-DEC-2001 (TREMELrel. 19, Created)
DT 01-DEC-2001 (TREMELrel. 19, Last sequence update)
DE 01-MAR-2003 (TREMELrel. 23, Last annotation update)
DE GHI3047P.
GN CG6114.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Berkely;
RA Stapleton M., Brokstein P., Hong L., Agbayani A., Carlson J.,
RA Champe M., Chavez C., Dorsett V., Farfan D., Frise E., George R.,
RA Gonzalez M., Guarin H., Li P., Liao G., Miranda A., Mungall C.J.,
RA Nunoo J., Pacht J., Paragae V., Park S., Phouanavong S., Wan K.,
RA Yu C., Lewis S.E., Rubin G.M., Celnik S.;
RL Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY060288; AAL25327.1; -.
DR FlyBase; FBgn0036544; CG6114.
DR InterPro; IPR000719; Prot_kinase.
DR Pfam; PF00069; pkinase; 1.
DR ProDom; PD000001; Prot_kinase; 1.
DR PROSITE; PS0011; PROTEIN_KINASE_DOM; 1.
KW ATP-binding; Transferase.
SQ SEQUENCE 701 AA; 76973 MW; 87B30879883B2425 CRC64;

Query Match 35.7%; Score 1248; DB 5; Length 701;
Best Local Similarity 44.9%; Pred. No. 2.6e-83;
Matches 284; Conservative 76; Mismatches 131; Indels 142; Gaps 19;

Qy 162 MASLQVGSLLTSCGSPHYACPEVIRGEKYDKRKYDVCVILFALLVAGLPDDNL 221
Db 1 MASLPAGSMLETS CGSPHYACPEVIRGEKYDKRKYDVCVILFALLVAGLPDDNL 60

Qy 222 ROLLEKVRGVPHFIPPPDCQSLLRGMIEVDAAARLTLEHIQHWY-IGGKNEPEPE 280
Db 61 ROLLEKVRGVPHFIPPPDCQSLLRGMIEVDAAARLTLEHIQHWY-IGGKNEPEPE 120

Qy 281 QPIPRKQVIRSLPSLEDIDPDVLDMSHISGCFDRDKLLDLSEENQEKWIYFLILDR 340
Db 121 LPMMEVQTHVPTATADPDVLDVNAISGCGKEKEKLIQELLSSHNTEKVIYFELLER 180

Qy 341 KERYPSQED--LPPRNE---IDPPKRVDSPLN----- 371
Db 181 KERRALEDDDDIAQKRSSELDAVPPKRLTCTRCINGTNPASYQISEGSLTPRROAF 240

Qy 372 -----RHKRRPERKSMVLSVTDGGSVPARRAEMAQ----- 405
Db 241 NFRSYSTRNHQRRSPPTVTSSVRS--SSYHSPTRCNSPSSAQQAANAISRPSAAGTR 299

Qy 406 ---HGORSISGASGLSTSLSS-----PRVTPHPSR---GSP--- 440
Db 300 HSTYGRDR--SGHSSVSRTPSHSQKSIEGVDVVVREPRERDSLRQEGGSGPRDR 357

Qy 441 --LPTPKGTP-----VHTKESPAGTNPPTP----- 464
Db 358 GDCGIPPGSGGSSGSGTSASPSVHRANSPTIALSMFHDPSNVNPNNGSPMWNSS 417

Qy 465 ----PSP--SVGVPMWRALNISKNSFLGSPRFRHKKIQQVTPPEMSNLTPESSPELAKK 519
Db 1 PPSPP--SVGVPMWRALNISKNSFLGSPRFRHKKIQQVTPPEMSNLTPESSPELAKK 519
```

```
Db 418 PCMPGSPCNTPGQLWKTRLTNIKNSFLGSPRFRHKKMQVSADE--VHLTPESSPELTGR 475
Qy 520 SWFGNFISLEKEQIFVVIKDKPLSSIKADIIVAFISLSLSHSVISOTSFRAYKATG- 578
Db 476 SWFGNLTTEKDETFTILVKGPATVAKAHLIHAFLSMAELSHSVSPTSFRVEYKRGNG 535
Qy 579 GRAVFOKPVKFOVDIT--YTEGEAQKENGIVSVTFTLLSGPSRRFRKRVVETIQALLST 636
Db 536 GPVNFQGHVKFQVDISAICKQGDIDM---LFTLTLLSGNIRFRICEHIQSQCCK 592
Qy 637 H-----DPP-----AAQHLSDTNCMEMMTGRLS 660
Db 593 RFPGPSPTVTGVTQAVSESSCGSVSERLS 625

RESULT 12
Q8SSX5 PRELIMINARY; PRT; 833 AA.
AC Q8SSX5;
DT 01-JUN-2002 (TREMELrel. 21, Created)
DT 01-JUN-2002 (TREMELrel. 21, Last sequence update)
DE 01-MAR-2003 (TREMELrel. 23, Last annotation update)
DE Putative serine/threonine protein kinase.
OS Dictyostelium discoideum (Slime mold).
OC Eukaryota; Mycetozoa; Dictyosteliida; Dictyostelium.
OX NCBI_TaxID=44689;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=AX4;
RA Gloeckner G., Eichinger L., Szafranski K., Pachebat J., Dear P.,
RA Lehmann R., Baumgart C., Parra G., April J.F., Guigo R., Kumpf K.,
RA Tungal B., Cox E., Quail M.A., Platzer M., Rosenthal A., Noegel A.A.;
RT "Sequence and Analysis of Chromosome 2 of Dictyostelium."
RL Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
DR EMBL; AC115685; AAL92711.1; -.
DR HSP; P24941; 1BUH.
DR InterPro; IPR000719; Prot_kinase.
DR InterPro; IPR002290; Ser_Thr_kinase.
DR Pfam; PF00069; pkinase; 1.
DR ProDom; PD000001; Prot_kinase; 1.
DR SMART; SM00220; S_TKC; 1.
DR PROSITE; PS0011; PROTEIN_KINASE_DOM; 1.
DR PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
KW ATP-binding; Kinase; Serine/threonine-protein kinase; Transferase.
SQ SEQUENCE 833 AA; 93467 MW; 51E64036C90C0F94 CRC64;

Query Match 24.9%; Score 870.5; DB 5; Length 833;
Best Local Similarity 31.1%; Pred. No. 2e-55;
Matches 215; Conservative 94; Mismatches 169; Indels 213; Gaps 16;

Qy 16 VGPVRLKTLKGGTGLVGLGVHCVTCQKVAIKIVNREKL--SSVLKMKVERETAILKLI 73
Db 6 VGPVRLKTLKGGTGLVGLGVHCVTCQKVAIKIVNREKL--SSVLKMKVERETAILKLI 65

Qy 74 EHPVLKLDVYENKYLVLVLEHVSQGLFDVLVKKGRLTTPKEARFQRIQISALDFCH 133
Db 66 DHPNWKMYEYETSKYFLILEVEGGLFDYLVEKGLSEGEALFFQIIGLEYCH 125

Qy 134 SHSTCHRDLPENLLLEKNNIRIADFGMASLQVGDLSLETSCGSPHYACPEVIRGEKYD 193
Db 126 SRNICHRLDLPENLLSCKRIKIADFGMGSIVRKMDLLHTSCGSPHYASPEVVSIGDYD 185

Qy 194 GRKADVSCGVILFALLVAGLPDDNRLQLLEKVRGVHFMHFIIPDCQSLLRGMIEV 253
Db 186 GQKADVSCGVILFALLVAGLPDDNRLQLLEKVRGVHFMHFIIPDCQSLLRGMIEV 245

Qy 254 DAARLTLEHIQHWYIGGKNEPEPEQIPRKVQIRSLFSLSDIDPDVLDMSHLSGCFR 313
Db 246 DPKSRISKEIKHPWFSNPNQKATPV-EEINAEPLVDYSQIDEDI FRSLLWALG-VG 303

Qy 314 DRNKLQDLSEENQEKWIYFLLDRKERY----- 344
Db 1 DRNKLQDLSEENQEKWIYFLLDRKERY----- 344
```

```
Db 304 TIDEVQQLVSNQKS--ATLIYYRLLEERKKFSDVNVKYGKPKETRRNSLSDMSLKIFPS 362
Qy 345 ----- 344
Db 363 GSNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 422
Qy 345 -----PSQEDDLPPRNBIDPPRKRV 365
Db 423 NNNNNNNNNNSAVGKSDSSSQPPHIQQPHSQQIPSNSTQESMQISPSNGAN--NMAI 480
Qy 366 DSPMLNRHGRKP-----ERKSMEVLSTVGGSPVAPARRAIEAQAHCQ 408
Db 481 QQPIINNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 536
Qy 409 RERSISG-----ASSGLSTSPSPRVTPHPSPRGSPLPKGTPTVHTPKE 454
Db 537 QQQQFNGNNNNNTMNVQPLTMSASSSSSSSTPSLSPNSTSTSTSPQLSAIKPDHY 596
Qy 455 SPAG--TPNPTPPSPSVGVGVPWRARLNSIKNSFLGSPFRHRKQVPP-----T 501
Db 597 QRRGSMTASTNPATSPTMS-----HRGKTSPTTEITSKVRKLKI 635
Qy 502 PEEMSLTPE-----SSPELAKSMFGNFIS 527
Db 636 SESQSN--TPNSPIIGSSP-----KKSMFSYFFS 662

RESULT 13
Q15524
ID Q15524 PRELIMINARY; PRT; 745 AA.
AC Q15524;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-FEB-1997 (TrEMBLrel. 02, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Serine/threonine protein kinase.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxId=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Colon;
RX MEDLINE=98399868; PubMed=9730619;
RA Espinosa L.; Navarro E.;
RT "Human serine/threonine protein kinase EMK1: genome structure and cDNA
cloning of isoforms produced by alternative splicing.";
RL Cytogenet. Cell Genet. 81:278-278(1998)
CC -1- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
DR EMBL; X97630; CAA66229.1; -
DR HSSP; Q63450; 1A06.
DR InterPro; IPR001772; Kinase Cterm.
DR InterPro; IPR000719; Prot_kinase.
DR InterPro; IPR002290; Ser_thr_kinase.
DR InterPro; IPR001245; Tyr_kinase.
DR InterPro; IPR000449; UBA_domain.
DR Pfam; PF02149; KAI; 1.
DR Pfam; PF00069; pkinase; 1.
DR Pfam; PF00627; UBA; 1.
DR PRINTS; PR00109; TYRKINASE.
DR ProDom; PD000001; Prot_kinase; 1.
DR SMART; SM00220; S_rkc; 1.
DR SMART; SM00165; UBA; 1.
DR PROSITE; PS00107; PROTEIN KINASE ATP; 1.
DR PROSITE; PS00111; PROTEIN KINASE DOM; 1.
DR PROSITE; PS00108; PROTEIN KINASE ST; 1.
KW ATP-binding; Kinase; Serine/threonine-protein kinase; Transferase.
SQ SEQUENCE 745 AA; 83204 MW; CEEA02EE762EC5DF CRC64;

Query Match 22.8%; Score 796.5; DB 4; Length 745;
Best Local Similarity 29.1%; Pred. No. 4.8e-50;
Matches 224; Conservative 120; Mismatches 243; Indels 183; Gaps 21;
Qy 10 AQHAQYGVPRUEKTLGKQGTGLVGLGVHCVTCQKVAIKVNRKLSVLMKVEREIAI 69
```

```
Db 11 ADEQPHIGNYELLKTIQKGNFAKVKLARHILTKGEVAVKIIDKTQLNSSSLQKLFREVRI 70
Qy 70 LKLTIEHPVLKLDHYENKLYLVLEHVSGGELFDYLKVGKRLTPKEARFPRIQIISAL 129
Db 71 MKVLNHPNIVKLFVETETKTLVLMVEYASGEVFDYLVVAHRMKKEARAKFRQIVSAV 130
Qy 130 DFCSHSICHRLDKPENLLDEKNIRIADFGMASLQVDSLLSTSCGSPHYACPEVIRG 189
Db 131 QYCHQKFLVHRDLKAENLLLDADNWKIADGFNEFTFGNKLDTFCGSPYAPAELEFG 190
Qy 190 EKYDGRKADVMSGCVIFPALLVGLPDDDDNLRLLEKVKGVFHMHPHFIIPDCOSLLRG 249
Db 191 KKYDGEVDVNSGLVILYTLVSGSLPFDQGNLKELRERVLGRKYRIPYMTSDCNLLKK 250
Qy 250 MIEVDAARLLEHIOKHIWYIGGKNBEPEPQIPRKVQIRSLPSLEDDIDPDVLDMSHL 309
Db 251 FLILNPSKRGTELEQIMKDRMNVGHEDDE-----LKPVEFLPDYK--DPRRTLMVSM 302
Qy 310 GCFRDRNKLLOLASEENQEKMIYFLLDRKERYPSOEDEDLPPRNEID----- 359
Db 303 CYTREE-----IQDSLVGQRYNEWMATYLLLYGYSSELEGDITLXPRPSADLTNSAQPPS 359
Qy 360 -----PPRKRVDS-----PMLNRHOK-----RRPE-----RKS----- 383
Db 360 HKVQSVSANPKQRRFSDQAGPAIPTSNYSKTKQSNNAENKRPEDRESGRKASSTAKV 419
Qy 384 -----EVLSTVDG-----GSPVPARRAIEMA--OHGQSRSSISG--A 416
Db 420 PASPLGLERKKTTPSTNTSVLSTNRSNPLERASLGQASIQNGKSLTPGSGRA 479
Qy 417 SSGLSLSTPLSPRVTPH-----PSPRGSLPTPKGTPTVHTPKESPA-- 457
Db 480 STASASAAVSAARQHQKMSASVHPNKAAGLPTESNCEVPRPSTAPQRPVVASPSAH 539
Qy 458 -----GTPNPT-----PSSPSVGVGVPWRARLNS 481
Db 540 NISSGGAPDRTNFPRGVSSRSTFHAGQLQRVDOQNLPGVTPASPSPGHSQGRGASGS 599
Qy 482 IKNSFLGSPFRHRKQVPP--TPEEMSLTP-----ESSPELAKSM--FG 523
Db 600 IFSKP--TSKVVRRNLNPEPKDRVETLRPHVSGGNDKKEEPREAKPSLRFTWSMK 657
Qy 524 NFISLEKEEFVWIKDKPLSSIKADIHAFSLPSLSHVSISQTSFRAEYKATGGPAVF 583
Db 658 TTSMENPEMREIRKVLNDANSQSELHEKYMLL--CMHGTPGHEDF----- 702
Qy 584 QKPVKQVDITVTEGGEAKENGISYVTFLLSGSPRRFRKRVETIQAO 633
Db 703 ---VQMEVEVC-----KLPRLSLNGVRFKRISGTSMAFKNIASKIANEL 743

RESULT 14
Q96HB3
ID Q96HB3 PRELIMINARY; PRT; 755 AA.
AC Q96HB3;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Similar to ELKI motif kinase.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxId=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Placenta;
RA Strausberg R.;
RL Submitted (MAY-2001) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
DR EMBL; BC008771; AA08771.1; -
DR HSSP; P24941; 1BUH.
DR Genew; HGNC:3332; MARK2.
```


GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 26, 2003, 11:50:16 ; Search time 10.9937 Seconds
(without alignments)
2857.432 Million cell updates/sec

Title: US-10-054-579-2

Perfect score: 3497

Sequence: 1 MTSTGKDGGAHQAVGYPV.....TNCNEMTGRLSKGIIPKS 668

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 127863 seqs, 47026705 residues

Total number of hits satisfying chosen parameters: 127863

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_41.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	790	22.6	1518	1 KKK1_YEAST	P34244 saccharomyc
2	785.5	22.5	512	1 K110_ARATH	Q38997 arabidopsis
3	779.5	22.3	752	1 MRK4_HUMAN	Q96134 homo sapien
4	769.5	22.0	774	1 MRK2_MOUSE	Q05512 mus musculu
5	764	21.8	576	1 SNF1_SCHPO	O74536 schizosacch
6	754	21.6	786	1 SNL1_HUMAN	P57059 homo sapien
7	749	21.4	776	1 SNL1_RAT	Q91455 rattus norv
8	742	21.2	776	1 MRK3_HUMAN	P27448 homo sapien
9	738	21.1	779	1 SNL1_MOUSE	Q60670 mus musculu
10	734	21.0	775	1 CDR2_SCHPO	P87050 schizosacch
11	729.5	20.9	552	1 AAK2_RAT	Q09137 rattus norv
12	729.5	20.9	915	1 KCC4_YEAST	P25389 saccharomyc
13	729	20.8	1142	1 GIN4_YEAST	Q12263 saccharomyc
14	726	20.8	619	1 SNF1_CANTR	O94168 candida tro
15	725.5	20.7	552	1 AAK2_HUMAN	P54646 homo sapien
16	723	20.7	548	1 AAK1_RAT	P54645 rattus norv
17	720.5	20.6	633	1 SNF1_YEAST	Q60782 saccharomyc
18	719	20.6	550	1 AAK1_HUMAN	Q13131 homo sapien
19	707	20.2	611	1 SNF1_CANGA	Q00372 candida gla
20	706	20.2	620	1 SNF1_CANAL	P52497 candida alb
21	688	19.7	502	1 RKI1_SECCB	Q02723 scale cere
22	647	18.5	891	1 KIN1_SCHPO	P22987 schizosacch
23	634	18.1	622	1 YNA3_CABEL	P45894 caenorhabdi
24	609.5	17.4	661	1 Y537_HUMAN	O60285 homo sapien
25	606	17.3	1064	1 KIN1_YEAST	P13185 saccharomyc
26	601	17.2	593	1 CDR1_SCHPO	P07334 schizosacch
27	558	16.0	800	1 KIN4_YEAST	Q01919 saccharomyc
28	551	15.8	1147	1 KIN2_YEAST	P13186 saccharomyc
29	522.5	14.9	714	1 HUNK_HUMAN	P57058 homo sapien
30	521	14.9	714	1 HUNK_MOUSE	O88866 mus musculu
31	504	14.4	353	1 ASK2_ARATH	P43292 arabidopsis
32	492.5	14.1	664	1 KCCB_HUMAN	Q13554 homo sapien
33	490	14.0	363	1 ASK1_ARATH	P43291 arabidopsis

RESULT 1
ID KKK1_YEAST STANDARD; PRT; 1518 AA.
AC P34244;
DT 01-FEB-1994 (Rel. 28, Created)
DT 01-FEB-1994 (Rel. 28, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE Probable serine/threonine-protein kinase YKL101W (EC 2.7.1.-).
GN YKL101W OR YKL453.
OS Saccharomyces cerevisiae (Baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
OX NCBI_Taxid:4932;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=S288c;
RX MEDLINE=94078677; PubMed=8256524;
RA Pallier C., Valens M., Puzos V., Fukuhara H., Cheret G., Sor F.,
RA Bolotin-Fukuhara M.;
RT "DNA sequence analysis of a 17 kb fragment of yeast chromosome XI
physically localizes the MRB1 gene and reveals eight new open reading
frames, including a homologue of the KIN1/KIN2 and SNF1 protein
kinases.";
RL Yeast 9:1149-1155(1993).
CC -!- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
CC NIM1 SUBFAMILY.

This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).

EMBL; X71133; CAA50456.1; --
EMBL; Z28101; CAA81941.1; --
PIR; S37928; S37928.
HSSP; Q63450; 1A06.
SGD; S0001584; HSL1.
GO; GO:0005935; C:bud neck; IDA.
GO; GO:0005940; C:septin ring; IDA.
GO; GO:0000086; P:G2/M transition of mitotic cell cycle; IGI.
GO; GO:0006468; P:protein amino acid phosphorylation; IDA.
GO; GO:0000074; P:regulation of cell cycle; IMP.
GO; GO:0000135; P:septin checkpoint; IGI.
InterPro; IPR000719; Prot_kinase.
InterPro; IPR002290; Ser_thr_kinase.
InterPro; IPR001245; Tyr_kinase.
Pfam; PF00069; pkinase; 1.
PRINTS; PR00109; TYRKINASE.
ProDom; PD000001; Prot_kinase; 1.
SMART; SM00220; S_TKc_1
PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
PROSITE; PS50011; PROTEIN_KINASE_DOW; 1.

ALIGNMENTS

KW Hypothetical protein; Transferase; Serine/threonine-protein kinase;
KW ATP-binding. 81 369 PROTEIN KINASE
FT NP BIND 87 95 ATP (BY SIMILARITY).
FT BINDING 110 110 ATP (BY SIMILARITY).
FT ACT SITE 239 239 BY SIMILARITY.
SQ SEQUENCE 1518 AA; 169592 MW; 803F84F7531241DD CRC64;

Query Match 22.6%; Score 790; DB 1; Length 1518;
Best Local Similarity 30.1%; Pred. No. 1.3e-26;
Matches 239; Conservative 107; Mismatches 32; Indels 182; Gaps 21;
Qy 16 VGPRLKTLKGGTGLVGLGVHCVTCQKVAIKVNRKL----- 55
Db 78 VGPWLKLTGKSSGSEVRLAKWETCOLAAIKVPKKAFVHCSSNGTVPNSSYSSMT 137
Qy 56 -----SSSVLMK-----VEREIALKLIIEHPHLKLDVVENKKYLVLYLHVHS 99
Db 138 SNVSSPSIASREHNSHNSQTPYGIERIEIVIMKLTISHTNVNVALFEVWENKSELYLVLEYVD 197
Qy 100 GGEFDYLVKKGRLTPKEARKFFROIISALDFCHSHSICHRLKPNLLIDKKN-NIRIA 158
Db 198 GGEFDYLVSKGLPEREAHYFKQIVGVSYCHSFNICHRLDKPENLLDKCKRRIKIA 257
Qy 159 DFGMASIQVGDLSLETSCGSPHYACPEVIRGEKYDGRKADVWSCGVLTFALLVGLPFD 218
Db 258 DFGWALELPLKLLKTSCTGSPHYASPIVNGRPHVCGSDVWSCGIVLTFALLTGHLPFD 317
Qy 219 DNLRLLEKVKRGVFMHPHFPDQCGLRGMEVDAAARLTLEHIOKHVIYGGKNEPE 278
Db 318 DNIKLLLVQSGKYQMPNSLSEARDLSKILVIDPEKRIITQELKH-PLIKKYDDL 375
Qy 279 PEQPIPEKQV-----IRSLPSLEDIPDVLDSMHSIGCFDRN 316
Db 376 PVNKLVRKMDKNWARKSNDLHLNVPSPVITLHSGKEIDSLRSQILWHGVGRE 435
Qy 317 KLLQDLISEENOKMYIFLLDLRKERY-----PSQDEDLPRNETDPPKRVDPMLN 371
Db 436 LIITAKLLQKPMSEKLYSLLOKQKHSISLSSSENKKSATESSVNEPIEVASKTAN 495
Qy 372 RHGKRPER-KMEVLVS-TDGGSPVPARRAE-----MAHQGRSRS----- 412
Db 496 NTGLRSENNDVKLHSLHSEIDTSTVNNNAITGVNTEINAPVLAQKQSFINTLSQPE 555
Qy 413 -----TSGASSG-----LSTSPSLSPRVTPHESPR 437
Db 556 SDKABEAVALPPIAIPFNASSRIFRNSYTSISRRSRRLSLNLSLSTASRSTVH 615
Qy 438 GSPLPKPGTPVHTPKSPAGTPNPTPPSPGVGVPWRARLNSIKNSFLGSPFRHRRKL 497
Db 616 DNEMLPQ-LPKSPSRYSLSRAIHASPSKSIH-----KSLSRKNIAATVAARLT 666
Qy 498 QVPTPEMSNLTPSSPELAKSHFGNFISLEKEQIFVVIKXPLSSIADIVHAFSLI 557
Db 667 Q-----NSASKRSLSLQSIKRS-----LNLDN-----LLVFDPLSPKPSASENVKSE 712
Qy 558 PSLSHVISOTSPPRA-----EYKATGPAVFOKPKVFOVDITYTE 597
Db 713 P-----HSLESDFEILCDQILFGNALDRILEEDNEKEKERTQORQNDTKSSADTTIS 769
Qy 598 GGEAKEN-----GIYSVFTLLSGPSRFRKV-----VETIOALLSTHDPAA 642
Db 770 GVSTNKENEGEYPTKIEKNFNMSYKPSENNWSGLSPFPIPEKENTLSSYLEBQPKRA 829
Qy 643 QHLSDTTNCMEM 655
Db 830 A-LSDDITNSFNKM 841

RESULT 2

K110_ARATH

ID K110_ARATH STANDARD; PRT; 512 AA.

AC Q38997; O04728; Q39076;

DT 15-JUL-1999 (Rel. 38, Created)
DT 15-JUL-1999 (Rel. 38, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE SNF1-related protein kinase KIN10 (RC 2.7.1.-) (AKIN10).
GN KIN10 OR SKIN10 OR AT3G01090 OR T4P13.22.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsids.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=93013041; PubMed=1339373;
RA le Guen L., Thomas M., Bianchi M., Halford N.G., Kreis M.;
RT "Structure and expression of a gene from Arabidopsis thaliana
RT encoding a protein related to SNF1 protein kinase.";
RL Gene 120:249-254(1992).
RN [2]
RP SEQUENCE FROM N.A.
RX STRAIN=cv. Columbia;
RA Lessard P., Kreis M., Thomas M.;
RT Submitted (DEC-1996) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RX STRAIN=cv. Columbia;
RL MEDLINE=21016720; PubMed=11130713;
RA Salanoubat M., Lemcke K., Bloeker H., Perez-Alonso M., Obermaier B.,
RA Partmann B., Valle G., Bloeker H., Rieger M., Ansoerge W., Unseid M.,
RA Deisenhoef V., Boutry M., Griwell L.A., Mache R., Puigdomenech P.,
RA De Simone V., Choise N., Artiguenave F., Robert C., Brottier P.,
RA Winkler P., Cattolico L., Weissenbach J., Saurin W., Quetier F.,
RA Schaefer M., Mueller-Auer S., Gabel C., Fuchs M., Benes V.,
RA Wurmbach E., Drzonek H., Erfle H., Jordan N., Bangert S.,
RA Wiedelmann R., Kranz H., Voss H., Holland R., Brandt P., Nyakatura G.,
RA Vezzi A., D'Angelo M., Pallavicini A., Toppo S., Simonati B.,
RA Conrad A., Hornischer K., Kauer G., Loehner T.-H., Nordstiek G.,
RA Reichelt J., Schaefer M., Schoen O., Barges M., Terol J., Climent J.,
RA Navarro P., Collado C., Perez-Perez A., Ottenwaelder B., Duchemin D.,
RA Cooke R., Laudie M., Berger-Llauro C., Purnelle B., Maury D.,
RA de Haan M., Maere A.C., Alcaraz J.-P., Cottet A., Casasuberta E.,
RA Monfort A., Argirou A., Flores M., Liguori R., Vitale D.,
RA Manhaupt G., Haase D., Schoof H., Rudd S., Zaccaria P., Mewes H.-W.,
RA Mayer K.F.X., Kaul S., Town C.D., Koo H.L., Tallon L.J., Jenkins J.,
RA Rooney T., Rizzo M., Waits A., Utterback T., Fujii C.Y., Shea T.P.,
RA Cressy T.H., Haas B., Maiti R., Wu D., Peterson J., Van Aken S.,
RA Pai G., Miltscher J., Sellers P., Gill J.E., Feldblyum T.V.,
RA Preuss D., Lin X., Nierman W.C., Salzberg S.L., White O., Venter J.C.,
RA Fraser C.M., Kaneko T., Nakamura Y., Sato S., Kato T., Asamizu E.,
RA Sasamoto S., Kimura T., Ideasa K., Kawashima K., Kishida Y.,
RA Kiyokawa C., Kohara M., Matsumoto M., Matsuno A., Muraki A.,
RA Nakayama S., Nakazaki N., Shinpo S., Takeuchi C., Wada T.,
RA Watanabe A., Yamada M., Yasuda M., Tabata S.;
RT "Sequence and analysis of chromosome 3 of the plant Arabidopsis
RT thaliana.";
RL Nature 408:820-822(2000).
RN [4]
RP SEQUENCE OF 1-19 FROM N.A.
RX STRAIN=cv. Columbia;
RX MEDLINE=95115691; PubMed=7816049;
RA le Guen L., Thomas M., Kreis M.;
RT "Gene density and organization in a small region of the Arabidopsis
RT thaliana genome.";
RL Mol. Gen. Genet. 245:390-396(1994).
CC -I- FUNCTION: MAY PLAY AN IMPORTANT ROLE IN A SIGNAL TRANSDUCTION IN
CC CASCADE REGULATING GENE EXPRESSION AND CARBOHYDRATE METABOLISM IN
CC HIGHER PLANTS.
CC -I- TISSUE SPECIFICITY: EXPRESSED IN ROOTS, SHOOTS AND LEAVES.
CC -I- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
CC SNF1 SUBFAMILY.
CC -I- SIMILARITY: Contains 1 UBA domain.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

```
CC EMBL; M93023; AAA32736.1; -
CC EMBL; X79707; CRA56146.1; -
CC EMBL; AC008261; AAF26165.1; -
CC EMBL; X94757; CRA64384.1; -
CC PIR; JCL146; JCL146.
CC HSSP; Q63450; IA06.
CC InterPro; IPR001772; Kinase Cterm.
CC InterPro; IPR000719; Prot Kinase.
CC InterPro; IPR002290; Ser Thr Kinase.
CC InterPro; IPR001245; Tyr Kinase.
CC InterPro; IPR000449; UBA_domain.
CC Pfam; PF02149; KAI; 1.
CC Pfam; PF00069; pkinase; 1.
CC Pfam; PF00627; UBA; 1.
CC PRINTS; PR00109; TYRKINASE.
CC ProDom; PD000001; Prot kinase; 1.
CC SMART; SM00220; STK; 1.
CC SMART; SM00165; UBA; 1.
CC PROSITE; PS00107; PROTEIN KINASE ATP; 1.
CC PROSITE; PS00111; PROTEIN KINASE DOM; 1.
CC PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
CC PROSITE; PS00030; UBA; 1.
CC Transferrase; Serine/threonine-protein kinase; ATP-binding;
KW Phosphorylation.
KW DOMAIN 19 271 PROTEIN KINASE.
FT DOMAIN 19 271 UBA.
FT NP_BIND 25 33 ATP (BY SIMILARITY).
FT BINDING 48 48 ATP (BY SIMILARITY).
FT ACT_SITE 142 142 BY SIMILARITY.
FT MOD_RES 175 175 PHOSPHORYLATION (AUTO-) (BY SIMILARITY).
SQ SEQUENCE 512 AA; 58373 MW; 5A18655A0AA506DF CRC64;
Query Match 22.5%; Score 785.5; DB 1; Length 512;
Best Local Similarity 41.9%; Pred. No. 6.3e-33;
Matches 160; Conservative 80; Mismatches 115; Indels 27; Gaps 7;
Qy 2 TSTKGGAQHAQYVGPYRLKTLGKGTGLVGLGHCVTCQKVAIKVNRKLSSEYL- 60
Db : : : : : : : : : : : : : : : : : : : : : : : : : :
4 SGTGSRSGVE--SILPNYKLGRTLGISGFRVKAIEHALTGKVAIKILNRKKNWEME 61
Qy 61 MKVRETAIKLIEHPVHLKLDVYENKYLVLVHVSQGLFDYLVKGRLLTPKARK 120
Db 62 EKVRREIKILRFVHPHPIRLYIEVTPTDYLVMVYNSGELFDYIVKGRLOQDEARN 121
Qy 121 FFRQIISALDPCGHSHICHRDLKPNILLDEKNIRIADFGMASIQVDSLETSCGSPH 180
Db 122 FFRQIISGVEYCHRMVVRDLKPNILLDSKCNVKTADFGLSNIMRDGHFLKTS CGSPN 181
Qy 181 YACEVIRGEKYDGRKADVWSCGVILFALLVGLPFDNDNLRLQLLEKVGKGVFMHPFIP 240
Db 182 YAAPEVISGKLYAGPEVDVWSCGVILVALLCGTLPPDDENIPNLFKKIKGIYTLPSHLS 241
Qy 241 PDCQSLLRGLMTEVDAARLTLEHQLKHWYGGKNEPEPEQIPRKVQI-----RSLPSLED 297
Db 242 PGARDLIPRLVUDVPMKRVVTIPEIRQHPWF-----QAHLPRLYLAVPDPVQOAKK 292
Qy 298 IDPDVLSMSHSLGCFRDRNKLQDLSEENQEKMIYFLLDLRKERYPS-----QEDE 350
Db 293 IDEILQEVNMGF--DENHLIESLRWNTQDGTYYLLDNRFRASSGYLGAEFQETM 350
Qy 351 DLPPRNEIDPPRKVDSPMLNR 372
Db 351 EGTGRMH---PAESVASPVSHR 369
```

RESULT 3

MRK4_HUMAN

```
MRK4_HUMAN STANDARD; PRT; 752 AA.
Q96L34; Q96JG7; Q9BYD8;
28-FEB-2003 (Rel. 41, Created)
28-FEB-2003 (Rel. 41, Last sequence update)
15-SEP-2003 (Rel. 42, Last annotation update)
MAP/microtubule affinity-regulating kinase 4 (EC 2.7.1.27)
(MAP/microtubule affinity-regulating kinase like 1).
MARK4 OR MARKL1 OR KIAA1860.
Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OC NCBI_TaxID=9606;
[1]
SEQUENCE FROM N.A. (ISOFORM 1).
TISSUE=Brain;
Drewes G., Mandelkow E.M.;
"MARK4, homologue of MARKL, MARK2 and MARK3.";
Submitted (SEP-2001) to the EMBL/GenBank/DBJ databases.
[2]
SEQUENCE FROM N.A. (ISOFORM 2).
TISSUE=Brain;
MEDLINE=21226021; PubMed=11326310;
Kato T., Satoh S., Okabe H., Kitahara O., Ono K., Kihara C.,
Tanaka T., Tsunoda T., Yamaoka Y., Nakamura Y., Furukawa Y.;
"Isolation of a novel human gene, MARKL1, homologous to MARK3 and its
involvement in hepatocellular carcinogenesis.";
Neoplasia 3:4-9(2001).
[3]
SEQUENCE FROM N.A. (ISOFORM 1).
TISSUE=Brain;
Beghini A., Magnani I., Roversi G., Piepoli T., DiTerlizzi S.,
Pollo B., Conti A.M.F., Cowell J.K., Finocchiaro G., Larizza L.;
"Neural progenitor-restricted isoform of MARKL1 gene is upregulated by
19q13 amplification in human glioblastoma.";
Submitted (JUN-2002) to the EMBL/GenBank/DBJ databases.
[4]
SEQUENCE FROM N.A. (ISOFORM 2).
TISSUE=Brain;
MEDLINE=21245130; PubMed=11347906;
Nagase T., Nakayama M., Nakajima D., Kikuno R., Ohara O.;
"Prediction of the coding sequences of unidentified human genes. XX.
The complete sequences of 100 new cDNA clones from brain which code
for large proteins in vitro.";
DNA Res. 8:85-95(2001).
[5]
SEQUENCE FROM N.A. (ISOFORM 1).
TISSUE=Thyroid;
Iscgai T., Ota T., Nishikawa T., Hayashi K., Otsuki T., Sugiyama T.,
Suzuki Y., Nagai K., Sugano S., Ishii S., Kawai-Hio Y., Saito K.,
Yamamoto J., Wakamatsu A., Nakamura Y., Kojima S., Nagahari K.,
Masuho Y., Ono T., Okano K., Yoshikawa Y., Aotsuka S., Sasaki N.,
Hattori A., Okumura K., Iwayanagi T., Ninomiya K.;
"NEDO human cDNA sequencing project.";
Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.
CC -/- CATALYTIC ACTIVITY: ATP + a protein = ADP + a phosphoprotein.
CC -/- ALTERNATIVE PRODUCTS:
CC Event=Alternative splicing; Named isoforms=2;
CC Name=1;
CC IsoId=Q96L34-1; Sequence=Displayed;
CC Name=2; Synonyms=MARKL1S;
CC IsoId=Q96L34-2; Sequence=VSP_004946;
CC -/- TISSUE SPECIFICITY: Ubiquitous. Isoform 2 is brain-specific.
CC -/- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES. MARK
SUBFAMILY.
CC -/- SIMILARITY: Contains 1 kinase-associated (KAI) domain.
CC -/- SIMILARITY: Contains 1 UBA domain.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL Outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (see http://www.isb-sib.ch/announce/

```


QY 539 KDKPLSS-----IKADIVHAFI-----SIPSLSHSVISOTSPRAEYKATGGPAVFQKP 586
 AC PRT: 776 AA.
 DB 578 SDTSITGLKAKAFQOLRTTKTGFLGKLNKXGLARQVQCAPARA-----SRGGLSPHAP 634
 QY 587 VKFQVDIITYTGGGAQKNGIYSVFTLLSGPSRRFRKRVETIQALLSTHDPAA 642
 DB 635 AQ-----SPGLHGAAGSREG-----WSLL-----BEVLE--QORLIQLQHPAA 672

RESULT 7

ID SNIL RAT STANDARD; PRT: 776 AA.
 AC Q9RIU5; Q9R081;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 15-SEP-2003 (Rel. 42, Last annotation update)
 DE Probable serine/threonine protein kinase SNF1LK (EC 2.7.1.1) (Salt-inducible protein kinase) (protein kinase KID2).
 GN SNF1LK OR SIK OR KID2.
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 OX NCBI_TaxID=10116;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Sprague-Dawley; TISSUE=Adrenal gland;
 RX MEDLINE=9330184; PubMed=1040390;
 RA Wang Z., Takemori H., Halder S.K., Nonaka Y., Okamoto M.;
 RT "Cloning of a novel kinase (SIK) of the SNF1/AMPK family from high salt diet-treated rat adrenal."
 RL FEBS Lett. 453:135-139(1999).
 [2]
 RP SEQUENCE FROM N.A.
 RA Feldman J.D., Vician L., Crispino M., Hoe W., Baudry M.,
 RA Herschman H.R.;
 RT "The KID2 gene encodes a protein kinase induced by depolarization in brain."
 RL Submitted (NOV-1998) to the EMBL/GenBank/DBJ databases.
 CC -!- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
 CC -!- SIMILARITY: Contains 1 UBA domain.

 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).
 CC -----

DR EMBL; AB020480; BA082673.1; --
 DR EMBL; AF106937; AAF14191.1; --
 DR HSBP; P24941; IAQ1.
 DR InterPro; IPR000719; Prot kinase.
 DR InterPro; IPR002290; Ser thr kinase.
 DR InterPro; IPR000449; UBA domain.
 DR Pfam; PF00069; pkinase; 1.
 DR ProDom; PD000001; Prot_kinase; 1.
 DR SMART; SM00220; S_TKc; 1.
 DR PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
 DR PROSITE; PS00111; PROTEIN_KINASE_DOM; 1.
 DR PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
 DR PROSITE; PS00030; UBA; 1.
 DR Transferrase; Serine/threonine-protein kinase; ATP-binding.
 KW DOMAIN 27 278
 FT DOMAIN 303 343
 FT NP_BIND 33 41
 FT BINDING 56 56
 FT ACT_SITE 149 149
 FT CONFLICT 473 473
 SQ SEQUENCE 776 AA; 84908 MW; 7BF745AF28F176E CRC64;

Query Match 21.4%; Score 749; DB 1; Length 776;
 Best Local Similarity 31.5%; Pred. No. 7e-31;
 Matches 214; Conservative 87; Mismatches 223; Indels 156; Gaps 20;
 QY 4 TGKGGGAQAQYVGYRLEKTKGQGLVKLGVCVTCOKVAIKIYNREKLSVLMKV 63
 DB 12 TGTGQGQQLRVLGVFVVERTLGKGNFAVVKLARHRTKTQVAIKIIDKTRLDSSNLEKI 71
 QY 64 ERETAIIKLIHHPVLKHDVYENKYLVLVLEHVSGLGDFVLVKKGLRTPKEARKEFR 123
 DB 72 YREVQLMKLNHPNIKLYQVMEFKOMLYIVTEFAKNGEMFDYLTNSGHLSENARKKF 131
 QY 124 QIISALDFCHSHSICHRLDKPENLLIDCKNIRIADFGMASLQVGDLSLTCSSGSPHYAC 183
 DB 132 QILSAVEYCHNHIVHDLAKTNLLDGNNDIKLADFGNFKYKPGEPSTWCGSPYAA 191
 QY 184 PEVIRGSKYGRKADVWCGVILFALLVGLALPDGDDNLRLEKVKRGVFMHPHFIPDQC 243
 DB 192 PEVFEKGEYEGPQIDVSLGVLVYVCGSLPDPGNLPTLRQVLEGRFRIPPFMSQDC 251
 QY 244 QSLRGMEVDAAARLTLEHIQKHIWYIGCKNPEP---EQIPRKVQIRSLPSLEDIDP 300
 DB 252 ETILRRMLVVDPAKRITIAIQIRHWM-----QADPTLQDDDPAFSMQGYTSNLGDYNE 306
 QY 301 DVLDSMHSILGCFDRNKLQDLSEENQEKMIYFLLDDR-----KERYPSQEDDLPPR- 355
 DB 307 QVLGIMQALGI--DRQRTVESLQNSSYNHFAIYLLERLREHRSQPSRATPAPARQ 364
 QY 356 -----NEI---DPPRKRVDSM-----LNRHGKRRP----- 378
 DB 365 POLRNSDLSSLEVPQEIPLPCDPRPSLLCPQALAQSLQAEIDCDLHSLQLPFPPLD 424
 QY 379 -----ERKSMEVLSVTGGSVPVARRAIEAQAQGRSISGASSGLST-----S 423
 DB 425 TNCSGVFRHRSISFSSLLDTAISEARQGPSLEEEQVQEPFGSTGRRHTLAEVSTHFS 484
 QY 424 PLSSPRVTPHSPRGSPLPTEPKGT-----VHPKESPA-----TPNPTPSSSPSVGG 472
 DB 485 PLNPPCIIVSSAAVSP---SEGTSSDCLPFSAEGPAGLGGGLATPGLLTSSP----- 537
 QY 473 VPMRRLNLSKNSFLGSPFRHRRKLQVTPPEMSN-----LTPESSE----- 515
 DB 538 -----VRLAS---PFLGS-----QSATPVLSQAGLGATVLPVPSFQECRRASDTSLT 582
 QY 516 -----LAKSWFGNFISLEKEQIFVVIKDKPLSSIKADIVHAFISPSLSHVSIS 566
 DB 583 QGLKAFQQRKQKARTKGFGLNKK-----IKGLARQV-C 615
 QY 567 QTSFRAEYKATGGPAVFQKP 586
 DB 616 QSSIRG---SRGGMSTFHTP 632

RESULT 8

ID MRK3 HUMAN STANDARD; PRT: 776 AA.
 AC P27438; O60219; Q8TE41; Q8W83; Q96RG1; Q9UMY9; Q9UN34;
 DT 01-AUG-1992 (Rel. 23, Created)
 DT 28-FEB-2003 (Rel. 41, Last sequence update)
 DT 15-SEP-2003 (Rel. 42, Last annotation update)
 DE MAP/microtubule affinity-regulating kinase 3 (EC 2.7.1.27) (Cdc25C-associated protein kinase 1) (CTAK1) (C-TAK1) (Serine/threonine protein kinase p78) (Ser/Thr protein kinase PAR-1) (Protein kinase SK10).
 DE MARK3 OR CTAK1.
 GN Homo sapiens (Human).
 OS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A. (ISOFORM 3).
 RX MEDLINE=98202387; PubMed=9543386;
 RA Peng C.Y., Graves P.R., Ogg S., Thoma R.S., Byrnes M.J. III, Wu Z.,

RA Stephenson M.T., Pivnicka-Worms H.;
RT "C-TAK1 protein kinase phosphorylates human Cdc25C on serine 216 and
RL promotes 14-3-3 protein binding."
RN Cell Growth Differ. 9:197-208(1998).
[2]
RP SEQUENCE FROM N.A. (ISOFORM 2).
RA Waggoner S.N., Stephen R., Farrar W.L., Howard O.M.Z.;
RT "Human serine/threonine protein kinase cTAK1/Kp78/Mark3;
RL Identification of a novel splice variant and a larger 5'UTR."
RN Submitted (JUN-1999) to the EMBL/GenBank/DBJ databases.
[3]
RP SEQUENCE FROM N.A. (ISOFORM 4).
RA MEDLINE=21326471; PubMed=11433294;
RA Sun T.Q., Lu B., Feng J.J., Reinhard C., Jan Y.N., Fantl W.J.,
RA Williams L.T.;
RT "FAR-1 is a Dishevelled-associated kinase and a positive regulator of
RL Wnt signalling."
RN Nat. Cell Biol. 3:628-636(2001).
[4]
RP SEQUENCE FROM N.A. (ISOFORM 5).
RA Drewes G.;
RT "Characterization of an alternatively spliced form of MARK3 from human
RL brain."
RN Submitted (JAN-2002) to the EMBL/GenBank/DBJ databases.
[5]
RP SEQUENCE FROM N.A. (ISOFORM 6).
RA Maheshwari K.K., Som S., Farra I.;
RL Submitted (JAN-1992) to the EMBL/GenBank/DBJ databases.
[6]
RP SEQUENCE FROM N.A. (ISOFORM 3).
RA TISSUE=Cervix;
RX MEDLINE=22388257; PubMed=12477932;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Heieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullany S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S.C., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smallos D.E.,
RA Schnerch A., Schein J.E., Jones S.J.M., Matra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length
RL human and mouse cDNA sequences."
RN Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
[7]
RP SEQUENCE OF 323-775 FROM N.A. (ISOFORM 7).
RA TISSUE=Urinary bladder;
RA Reynolds C.H., Patel U.A., Anderton B.H.;
RT "Homo sapiens mRNA partial sequence for a protein kinase, STK10,
RL similar to p78/C-TAK1."
RN Submitted (JUL-1999) to the EMBL/GenBank/DBJ databases.
CC -1- FUNCTION: Involved in the specific phosphorylation of microtubule-
associated proteins for tau, MAP2 and MAP4. Phosphorylated Cdc25C
on Ser-216.
CC -1- CATALYTIC ACTIVITY: ATP + a protein = ADP + a phosphorylated protein.
CC -1- ALTERNATIVE PRODUCTS:
CC Event=Alternative splicing; Named isoforms=7;
CC Name=1;
CC IsoId=P27448-1; Sequence=Displayed;
CC Name=2; Synonyms=CTAK75a;
CC IsoId=P27448-2; Sequence=VSP_004944;
CC Name=3;
CC IsoId=P27448-3; Sequence=VSP_004942, VSP_004944;
CC Name=4;
CC IsoId=P27448-4; Sequence=VSP_004942, VSP_004945;
CC

CC Name=5;
CC IsoId=P27448-5; Sequence=VSP_004942;
CC Name=6; Synonyms=PS8;
CC IsoId=P27448-6; Sequence=VSP_004942, VSP_004943, VSP_004944;
CC Name=7;
CC IsoId=P27448-7; Sequence=VSP_004943;
CC -1- TISSUE SPECIFICITY: Ubiquitous
CC -1- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES. MARK
CC SUBFAMILY.
CC -1- SIMILARITY: Contains 1 kinase-associated (KAI) domain.
CC -1- SIMILARITY: Contains 1 UBA domain.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See [http://www.isb-sib.ch/](http://www.isb-sib.ch/announce/)
or send an email to license@isb-sib.ch).
CC EMBL; U64205; AAC15093.1; -;
CC EMBL; AF159295; AAD48007.1; -;
CC EMBL; AF387637; AAK82367.1; -;
CC EMBL; AF465413; AAL69982.1; -;
CC EMBL; M80359; AAA59991.1; -;
CC EMBL; BC024773; AAH24773.1; -;
CC EMBL; AF170723; AAD51631.1; -;
CC HSSP; Q63450; IA06.
CC Genew; HGNC:6897; MARK3.
CC MIM; 602678; -;
CC InterPro; IPR001772; Kinase Cterm.
CC InterPro; IPR000719; Prot_kinase.
CC InterPro; IPR002290; Ser_thr_kinase.
CC InterPro; IPR001245; Tyr_kinase.
CC InterPro; IPR000449; UBA_domain.
CC Pfam; PF02149; KAI; 1.
CC Pfam; PF00069; pkinase; 1.
CC Pfam; PF0627; UBA; 1.
CC PRINTS; PR00109; TYRKINASE.
CC ProDom; PD000001; Prot_kinase; 1.
CC SMART; SM00220; S_TKc; 1.
CC SMART; SM00165; UBA; 1.
CC PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
CC PROSITE; PS00111; PROTEIN_KINASE_DOM; 1.
CC PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
CC PROSITE; PS00303; UBA; 1.
CC Transferase; Serine/threonine-protein kinase; ATP-binding;
KW Alternative splicing.
FT DOMAIN 56 330 PROTEIN KINASE.
FT DOMAIN 326 365 UBA.
FT DOMAIN 727 776 KAI.
FT NP_BIND 62 70 ATP (BY SIMILARITY).
FT BINDING 85 85 ATP (BY SIMILARITY).
FT ACT_SITE 201 201 BY SIMILARITY.
FT VARSPLIC 162 184 Missing (in isoform 3, isoform 4, isoform
5 and isoform 6).
FT VARSPLIC 394 409 /FTId=VSP_004942.
FT VARSPLIC 638 661 Missing (in isoform 6 and isoform 7).
FT VARSPLIC 638 646 Missing (in isoform 2, isoform 3 and
isoform 6).
FT VARSPLIC 638 646 /FTId=VSP_004944.
FT CONFLICT 125 125 Missing (in isoform 4).
FT CONFLICT 139 139 E -> Q (IN REF. 5).
FT CONFLICT 149 149 E -> K (IN REF. 4 AND 5).
FT CONFLICT 448 448 R -> K (IN REF. 2).
FT CONFLICT 466 466 A -> G (IN REF. 5).
FT CONFLICT 479 479 S -> G (IN REF. 1 AND 2).
FT CONFLICT 479 479 S -> T (IN REF. 5).
FT CONFLICT 539 539 A -> D (IN REF. 5).
FT CONFLICT 626 626 S -> T (IN REF. 5).
FT CONFLICT 668 668 E -> K (IN REF. 5).

SEQ	SEQUENCE	776 AA; 86944 MW; A245496849070098 CRC64;	21.2%; Score 742; DB 1; Length 776;
Query Match	Best Local Similarity	29.0%; Pred. No. 1.6e-30;	
Matches	218; Conservative	114; Mismatches	269; Indels 150; Gaps 20;
QY	10	AQAQYVGPRLKTLKGGQTGLVGLVCHVTCQKVAIKIWNREKLSEVLMKVREIAI	69
DB	47	ADBPQHIGNYRLKTIKGNFAKVKLARHLITGREVAIKIITDKLTQNLPTSKQLFREVI	106
QY	70	LKLEIHPHLKLDHVDYENKKYLYLVLEHVSGLGLFDYLVVKKGRITPKEARFFRQ	124
DB	107	MKILNHPNIVKLFEVITEKTYLIMEVAGGEVFDYLVAGRKWEKARSKFGQCGAQ	166
QY	125	-----IISALDFCHSHSTCHRDLKPENLLLDKNNRIADFGMASLQ	166
DB	167	QTIKVQVSDLLSLMFTFIYSAVQYCHQKRIVHRDLKAENLLLDADNMNIADFGSNEF	226
QY	167	VGDSILETSCGSPHYACPEVIRGSKYDGRKADWCSGVILPALLVGAALPFDDNLRLQLE	226
DB	227	TVGQKLDTFCGSPPYAAPLPELFGQKKYDGPEDVWMSGLVILTVSGSLPFPDQNLKELRE	286
QY	227	KVRKGVFHMFHPTPCDCQSLRGLMIEVDAAERLLTLEHIQKHIWIGKNKEPEPQTPRK	286
DB	287	RVRGKYRIIFYMSTDCENLLKRLFLVNLPIKRGTLQIOMKDRWINAGHEEDLPFVE--	344
QY	287	VQIRSLPSLEIDIPDVLDSMHSLSGCFDRNKLQDLLSEENQKMTYFLLLDKRYPS	346
DB	345	-----PELDISQKRIDIWGMGYSQEE---IQESLSKKYDITATYLLGRKS---S	392
QY	347	QED-EDLPPNEIDPPKRVDSPLNHRGKRRPERKMSMEVLSVTD-----GGSPVPA	397
DB	393	ELDASDSSSSNLSLAKVRPSSDLNNGTG-QSPHHKVRQSVSSQKORRYSDHAGPAIPS	451
QY	398	RRAI-----ENAQGQRSISGASHSLSTSPLSPRVTPHPSPRGSPLEPTPKG	446
DB	452	VVAYPEKRSQTADSLKEDPISSRKSSGSAVGGKGIAPASPLGNASNPNKADIPERKK	511
QY	447	TPVHTPKESPA-----GTNPPTPPS-----SPSVG	471
DB	512	SST-VPSSNTASGGMTRNTYVCSERTADRHVSIVQNGKENTIPDORTPVASTHSSA	570
QY	472	GVPRARL-----NSIKNSFLGSPRFRRLQVPTPEMSNLTPESSPELAKKS-----WFG	523
DB	571	ATPDRIFRPGTASRSTFHGQPR-ERRATVYGGPPASPSSLSHEATPLSQTSRGSTNLS	629
QY	524	NFTS-LEKEEQIFVVIKDKPLSSIKADIVHAFISPLSHSVISQTS-----F	570
DB	630	KLTSKLTNRNMSFRFKRLRTEYERNG-----RYEGSSRNVSVAQDENKEAKPRSLRF	683
QY	571	RABYKATGG--PAVFQKPVKQVD-----ITYTEGGBAQKEN-----	605
DB	684	TWGMKTTSSMDPDGMREIRKVLDAANNCDYEQRRERFLFCVHGUGHAEINLVQMEVCKL	743
QY	606	---GIYSVTFLLSGSPRRFRKRVVETIQAQL	633
DB	744	PRLSLNGVRFKRISGTSIAFNKIASKIANEL	774
RESULT 9			
SNIL MOUSE			
AC	ID	SNIL MOUSE	STANDARD; PRT; 779 AA.
AD	AC	Q60670;	
DT	DT	01-NOV-1997 (Rel. 35, Created)	
DT	DT	16-OCT-2001 (Rel. 40, Last sequence update)	
DT	DT	15-SEP-2003 (Rel. 42, Last annotation update)	
DE	DE	Probable serine/threonine protein kinase SNF1LK (BC 2.7.1.-) (HRT-20)	
DE	DE	(Myocardial SNF1-like kinase).	
GN	GN	SNF1LK OR MSK.	
OS	OS	Mus musculus (Mouse).	
OC	OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;	
OC	OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.	
NCBI	NCBI	TaxID=10090;	

```
[1] SEQUENCE FROM N.A., AND REVISIONS TO 8; 16-18; 44-45; 316 AND 435.
```

```
RP TISSUE=Embryo;
```

```
RZ Ruiz J.C.;
```

```
RA Submitted (Oct-1999) to the EMBL/GenBank/DDBJ databases.
```

```
RN [2]
```

```
RC SEQUENCE OF I-435 FROM N.A., AND TISSUE SPECIFICITY.
```

```
RD TISSUE=Embryo;
```

```
RE MEDLINE=95200798; PubMed=7893599;
```

```
RF Ruiz J.C., Conlon P.L., Robertson E.J.;
```

```
RG "Identification of novel protein kinases expressed in the myocardium
```

```
RL of the developing mouse heart.";
```

```
RT Mech. Dev. 48:153-164(1994).
```

```
CC CC -I- TISSUE SPECIFICITY: EXPRESSED IN LUNG, SKIN, OVARY, HEART AND
```

```
STOMACH. NO EXRESSION IN BRAIN, LIVER OR SKELETAL MUSCLE.
```

```
CC CC -I- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
```

```
CC SNF1 SUBFAMILY.
```

```
CC CC -I- SIMILARITY: Contains 1 UBA domain.
```

```
-----
```

```
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
```

```
between the Swiss Institute of Bioinformatics and the EMBL Outstation at
```

```
the European Bioinformatics Institute. There are no restrictions on its
```

```
use by non-profit institutions as long as its content is in no way
```

```
modified and this statement is not removed. Usage by and for commercial
```

```
entities requires a license agreement (See http://www.isb-sib.ch/announce/ or send an email to licensee@isb-sib.ch).
```

```
-----
```

```
EMBL; UI1494; AAA67926.2; --
```

```
HSP; P24941; IAQL.
```

```
MGD; MGII:104754; Snfllk.
```

```
DInterPro; IPRO00719; Prot kinase.
```

```
IInterPro; IPR002290; Ser thr pkinase.
```

```
Pfam; PF00069; pkinese.i.
```

```
ProDom; PD000001; Prot_kinase; 1.
```

```
SMART; SM00220; S_TKc; 1.
```

```
PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
```

```
PROSITE; PSS0011; PROTEIN_KINASE_DOM; 1.
```

```
PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
```

```
DR PROSITE; PSS0030; UBA; 1.
```

```
KW Transferase; Serine/threonine-protein kinase; ATP-binding.
```

```
FT DOMAIN 27 278 PROTEIN KINASE.
```

```
FT BINDING 303 343 UBA.
```

```
FT NP_BIND 33 41 ATP (BY SIMILARITY).
```

```
FT BINDING 56 56 ATP (BY SIMILARITY).
```

```
FT ACT_SITE 149 149 BY SIMILARITY.
```

```
SQ SEQUENCE 779 AA; 85027 MW; 7B0813BC46D9C4E CRC64;
```

```
Query Match      21.1%; Score 738; DB 1; Length 779;
```

```
Best Local Similarity 33.9%; Pred.No. 2.5e+30;
```

```
Matches 189; Conservative 83; Mismatches 193; Indels 92; Gaps 14;
```

```
Qy   4 TGKGGAQAHAQYGPVRLEKTIGQGTLVKLGVCVCTOKVAIKIVNREKLSEVLMKV 63  
    :|::||::||::||::||::||::||::||::||::||::||::||::||:  
Db   12 SGTGCGQQPKPLRVGVDFVERTLKGNFAVVCLARHVRTKTQVAIKIIDTRLDSSNLEKI 71
```

```
Qy   64 ERRIATLIKLEHPHLKHDDVVENKKYLVIHLVHSVGGLFDLVKKGRLTPKEARKFFR 123  
    ||::||::||::||::||::||::||::||::||::||::||::||:  
Db   72 YREVQMKNLNHNII::IKYVMETMDMLIVTFNAKNGMFIDLTNSGHLSENEAREQKFV 131
```

```
Qy   124 QITISALDFCHSHSICHRLDKPENLLDEKNNIRIAIDFGMASLQVGSLLSETSCSGSPHYAC 183  
    ::||::||::||::||::||::||::||::||::||::||::||:  
Db   132 QLISAVEYCNHHIVHRDLKTENLLDSNMIDLADFGFCNFYPKPGEPISTCVGPSPPYAA 191
```

```
Qy   184 PEVIRGEKYDGRRADVWSCGVILFALLVAGLPEDDDLNLROLLEBKVKRGVFHMPHIPDPDC 243  
    ||::||::||::||::||::||::||::||::||::||::||::||:  
Db   192 PEVEPEGKEYEGPDVMSLGWLVIYLCVCSLPDPGNPLPTLRQRVLGRFRIFPFMFSQCDC 251
```

```
Qy   244 QSLLRGNIENVDAARRLTLEHQHIWYGCKNEPEP---EQTPRKVQTSLPSLEDIDP 300  
    ::||::||::||::||::||::||::||::||::||::||::||:  
Db   252 ETIRRMVLVDPAKRITTAIQIRHRMW----QADPTLIQQDDPAFDMOGYTSLNLGDYNE 306
```

```
Qy   301 DVLSDMHSICGFDRNKILLDLLISEENOEKMTYLFLLDRKERY----PSOEDELAPPR-- 355
```

```
Db 307 QVLGIMQALGI--DRQRTIESLQSSYNHFAAIIYLLERLKEHRSQAFSSRRTPAPTRQ 364
Qy 356 -----NET---DPERKRVDSPM-----LNRHGKRP----- 378
Db 365 PQLRSSLSLEVPQELPCDPRFRPSLCPQOPALQSVLQAIQDCLHSLQPLFLPLD 424
Qy 379 -----ERKSMEVLSYTDGGSVPARRAJEMAHQGRSRSISGASSGLST-----S 423
Db 425 TNCSGVFRHRSISPSLLDTAISERARQGPFSLEEEQVQEPPLGSGRRHTLAEVSTHFS 484
Qy 424 PLASPRVTPHPSRGSLPTPKGTP-----VHPKESPAQ-----TPNPTPSSSVSG 472
Db 485 PLNPPCCIIVSSATASP---SEGTSDDCLPFSASEGPAGLGSLATPGLLTSSP----- 537
Qy 473 VPMRRLNLSIKNSPLGS 489
Db 538 -----VRLAS---PPLGS 547

RESULT 10
CDR2 SCHPO STANDARD; PRT; 775 AA.
AC P87050;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Mitosis inducer protein kinase cdr2 (EC 2.7.1.-).
GN CDR2 OR SPAC57A10.02.
OS Schizosaccharomyces pombe (fission yeast).
OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
OC Schizosaccharomycetales; Schizosaccharomycetaceae;
OC Schizosaccharomycetes;
OX NCBI_TaxID:4896;
RN [1]
RP SEQUENCE FROM N.A., AND FUNCTION.
RC STRAIN:972;
RA Breeding C.S., Hudson J., Balasubramanian M.K., Hemmingen S.M.,
RA Young P.G., Gould K.L.;
RT "The cdr2(+)" gene encodes a regulator of G2/M progression and
RT cytokinesis in Schizosaccharomycetes pombe.";
RL Mol. Biol. Cell 9:3399-3415(1998).
[2]
RN SEQUENCE FROM N.A.
RP STRAIN:972;
RX MEDLINE=99060136; PubMed=9843577;
RA Breeding C.S., Hudson J., Balasubramanian M.K., Hemmingen S.M.,
RA Young P.G., Gould K.L.;
RA "The cdr2(+)" gene encodes a regulator of G2/M progression and
RA cytokinesis in Schizosaccharomycetes pombe.";
RL Mol. Biol. Cell 9:3399-3415(1998).
[2]
RN SEQUENCE FROM N.A.
RP STRAIN:972;
RX MEDLINE=21848401; PubMed=11859360;
RA Wood V., Gwilliam R., Rajandream M.A., Lyne M., Lyne R., Stewart A.,
RA Sgouras J., Peat N., Hayles J., Baker S., Basham D., Bowman S.,
RA Brooks K., Brown D., Brown S., Chillingworth T., Churcher C.M.,
RA Collins M., Connor R., Cronin N., Davis P., Feltwell T., Fraser A.,
RA Gentles S., Goble A., Hamlin N., Harris D., Hidalgo J., Hodgson G.,
RA Holroyd S., Hornsby T., Howarth S., Huckle E.J., Hunt S., Jagels K.,
RA James K., Jones L., Jones M., Leather S., McDonald S., McLean J.,
RA Mooney P., Moulé S., Mungall K., Murphy L., Niblett D., Odell C.,
RA Oliver K., O'Neill S., Pearson D., Quail M.A., Rabinowitch E.,
RA Rutherford K., Rutter S., Saunders R., Seeger K., Sharp S.,
RA Skelton J., Simmonds M., Squares R., Squares S., Stevens K.,
RA Taylor K., Taylor R.G., Tivey A., Walsh S.V., Warren T., Whitehead S.,
RA Woodward J., Volkheart G., Aert B., Robben J., Grynoprez B.,
RA Weltjens I., Vanstreels E., Rieger M., Schaefer M., Mueller-Auer S.,
RA Gabel C., Fuchs M., Fritz C., Holzer E., Moestl D., Hilbert H.,
RA Borzym K., Langer I., Beck A., Leirach H., Reinhardt R., Pohl T.M.,
RA Eger P., Zimmermann W., Wedler H., Wambutt R., Purnelle B.,
RA Goffeau A., Cadieu E., Dreano S., Gloux S., Lelaure V., Mottier S.,
RA Galibert F., Aves S.J., Xiang Z., Hunt C., Moore K., Hurst S.M.,
RA Lucas M., Rochet M., Gaillardin C., Tallada V.A., Garzon A., Thode G.,
RA Daga R.R., Cruzado L., Jimenez J., Sanchez M., del Rey F., Benito J.,
RA Dominguez A., Revuelta J.L., Moreno S., Armstrong J., Forsburg S.L.,
RA Cerrutti L., Lowe T., McCombie W.R., Paulsen I., Potashkin J.,
RA Shpakovski G.V., Ussery D., Barrell B.G., Nurse P.;
RT "The genome sequence of Schizosaccharomycetes pombe.";
RL Nature 415:871-880(2002).
```

```

(3)
RN MEDLINE=99060131; PubMed=9843572;
RX Kanoh J., Russell P.;
RT "The protein kinase Cdr2, related to Nimi/Cdr1 mitotic inducer,
RT regulates the onset of mitosis in fission yeast.";
RL Mol. Biol. Cell 9:3331-3334(1998).
CC !- FUNCTION: Acts as a mitotic inducer. In G2 it negatively regulates
CC weel, a mitotic inhibitor. Also has a role in cytokinesis where it
CC is required for proper septum formation.
CC !- PTM: Autophosphorylated.
CC !- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
CC NIMI SUBFAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; Z94864; CAB08165.1; --
CC EMBL; AF092508; AAC72832.1; --
CC FIR; T38929; T38929.
CC HSSP; Q63450; 1A06.
CC GeneDB_SPombe; SPAC57A10.02; --
CC InterPro; IPR000719; Prot kinase.
CC InterPro; IPR002290; Ser Thr kinase.
CC Pfam; PF00069; pkinase; 1.
CC ProDom; PD000001; Prot kinase; 1.
CC SMART; SM00220; S_TK; 1.
CC PROSITE; PS00107; PROTEIN KINASE ATP; FALSE_NEG.
CC PROSITE; PS00111; PROTEIN KINASE DOM; 1.
CC PROSITE; PS00108; PROTEIN KINASE_ST; 1.
CC Transferase; Serine/threonine-protein kinase; ATP-binding;
CC Mitosis; Phosphorylation.
CC DOMAIN 10 262 PROTEIN KINASE (BY SIMILARITY).
CC NP_BIND 16 24 ATP (BY SIMILARITY).
CC ACT_SITE 133 133 BY SIMILARITY.
CC FT BINDING 39 39 ATP (BY SIMILARITY).
CC SEQUENCE 775 AA; 85971 MW; 2E2EECABC8A4FAC CRC64;

Query Match 21.0%; Score 734; DB 1; Length 775;
Best Local Similarity 29.6%; Pred. No. 4e-30;
Matches 228; Conservative 105; Mismatches 234; Indels 204; Gaps 27;

Qy 16 VGPYRLKTLGKGQTGLVGLGVHCVTCQVAK-IVNREKLSVLMKVERETAILKLE 74
Db 7 VGPWEGLSLGSGGPNSSRLAKHRETGQLAVKPIVGNSELSSQQAIEGELVLRLE 66
Qy 75 HPVLKLDVYENKLYLVLEHVSQGEFLDYLVKGRLLTPKEARKFFRQIISALDFCHS 134
Db 67 HPNVQLIDVISAQEQLFVVVEYMPGGEFLDCMLRKGSTEQDTAKFLWQLGLEVYCHK 126
Qy 135 HSTCHRDLPENLLDEKNIRIADPGMASLQVDSLLTSCGSPHYACPEVIRGEKYDG 194
Db 127 LHCHRDLPENLYLDAHSGIKIGFEFGMASIQQPGKLLQTCGSPHYASPEIIMGSRSDG 186
Qy 195 RKADVMSGVILFALLVAGLPPDDNRLRLLEKVKGVFMHPHFIPDCQSLRGMIEVD 254
Db 187 CASDIWSCGLIIFALTGTGLPDDNNIRSLLLKVCQGFEMFSNPQAQHLIYRLMDVD 246
Qy 255 AARLTLEHIQHIWYIGGKNRP-----EPQPIPRKQVIRSLPSLEIDDPVLSMHS 308
Db 247 SSTRTITMEQIREH-----PFLSCFVHPNISIP-----IISAP-IQPIDPLIVQHLSL 292
Qy 309 L-GCFRDRNKLQDLISEENQEKMYIFLLDRKERYPSQEDLPPEINIDPRKRVDS 367
Db 293 VFRCSDDPPLYEKLASQSPLEKTYLTLL--SRHLHP-----PSSAAVDNRRAVDD 343
Qy 368 PM--LNRHGKRRPERKSMELSVLTDGGSVPARRAJEMAHQGRSRS-----ISGA-- 416
```



```
FT DOMAIN 17 29 POLY-HIS.
FT NP_BIND 52 303 PROTEIN KINASE.
FT BINDING 58 66 ATP (BY SIMILARITY).
FT ACT_SITE 81 81 ATP (BY SIMILARITY).
FT MOD_RES 174 174 BY SIMILARITY.
SQ SEQUENCE 207 207 PHOSPHORYLATION (AUTO-) (BY SIMILARITY).
SEQUENCE 619 AA; 70323 MW; 0FCF1FC3DCE706D7 CRC64;

Query Match 20.8%; Score 726; DB 1; Length 619;
Best Local Similarity 33.1%; Pred. No. 7.8e-30;
Matches 173; Conservative 105; Mismatches 168; Indels 76; Gaps 15;

Qy 7 DGAQAQYVGPYRLEKTLGKQGLVGLVCHVCTCKVAIKIVNRKLSVLM-KVER 65
Db 40 DPNVNPANRIGRYQIIKTLGEGSGFKVLAHQHVGTKVALKIINRKTAKSDMQGRVER 99
Qy 66 EIALKILIEHPHVLKHDVYENKKYLVLVLEHVSGBELFDYLVKKGRLTPKEARKFRQI 125
Db 100 EISYLRLLRPHIILKLDYVKSDEIIMVIE-FAGKELFDYVQGRKMPDEARRFFQOI 158
Qy 126 ISALDFCHSHSICHRLDKPENLLLDKNNIRIADFGWASLQVGDLSLETSCGSPHYACPE 185
Db 159 IAAVEYCHRHKIVHRDLKPNLLDDQNLNVIADFGLSNMTDGNFLKTSFGSPNYAAPE 218
Qy 186 VIRGEKIDGRKADYVWCGVILFALLVGLALPDDDLRLQLLEKVKRGVFMHPHFPDPCQS 245
Db 219 VTSCKLYAGPEVDWSSGVILVYMLCGRLPDDEFIPALFKINSYVYTLPNYLSFGAKH 278
Qy 246 LLRMQIEVDAAARLTLLEHQHVIWIGKNEPPEQIP---RKQVTRSLPSLEDDIPDV 302
Db 279 LLTRMLVNPVNLRTIHEIMEDEF---KQDMPDYLLPDLKSKIKTSKI-----DIDEDV 330
Qy 303 LDSMHSIGCFRDRNKLQDLASEENQKMIYFLLDRKERYPSQEDLPPRNEIDPPR 362
Db 331 ISAL-SVTMGYDREIIS--VIEKANREAA----- 358
Qy 363 KRVDSPMLNRHGRKRPKRSMEVLSTVDGGSPVPARRAIEAQAQGRSRSISGASSGLST 422
Db 359 -----GGATPTNQSKSTNEVLDAVLLMKENHTLVKDLKKSSENI---ESFLSL 404
Qy 423 SPLSPRVTPHPSRGSPLTPKGTVPHTPKESFAGTNPPTPPSPSPVGGVFWEARLNSI 482
Db 405 SP---PPSSSPNP-GSTSSAP-GVQOSLTYQTTLATVPDLSTLPSNSTAILP---TSLPSI 457
Qy 483 KNSFLGSPFRHRRKLVQVTPPEMSNLTPESSPELAKKSW-FG 523
Db 458 HRAYMTEKVNDDPQQQIPAPQPTKKL-----KTRWHFG 490

RESULT 15
AAK2 HUMAN
AC P54646; Q9H188; O9UD43, PRT; 552 AA.
DT 01-OCT-1996 (Rel. 34, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE 5'-AMP-activated protein kinase, catalytic alpha-2 chain (EC 2.7.1.-) (AMPK alpha-2 chain).
GN PRKAA2 OR AMPK2 OR AMPK.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI TaxID=9606;
[1] _SEQUENCE FROM N.A.
RP TISSUE=Heart;
RC MEDLINE=95047501; PubMed=7959015;
RA Aguan K., Scott J., See C.G., Sarkar N.H.;
RT "Characterization and chromosomal localization of the human homologue
RT of a rat AMP-activated protein kinase-encoding gene: a major
RT regulator of lipid metabolism in mammals.";
RL Gene 149:345-350(1994).
RN
```

```
RP SEQUENCE FROM N.A.
RC TISSUE=Skeletal muscle;
RX MEDLINE=95080410; PubMed=7988703;
RA Beri R.K., Marley A.E., See C.G., Sopwith W.F., Aguan K., Carling D.,
RA Scott J., Carey F.;
RT "Molecular cloning, expression and chromosomal localisation of human
RT AMP-activated protein kinase.";
RL FEBS Lett. 356:117-121(1994).
RN [3]
RP SEQUENCE OF 33-552 FROM N.A.
RA Cobley V.;
RL Submitted (APR-2000) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: RESPONSIBLE FOR THE REGULATION OF FATTY ACID SYNTHESIS
CC BY PHOSPHORYLATION OF ACETYL-COA CARBOXYLASE. IT ALSO REGULATES
CC CHOLESTEROL SYNTHESIS VIA PHOSPHORYLATION AND INACTIVATION OF
CC HORMONE-SENSITIVE LIPASE AND HYDROXYMETHYLGUTARYL-COA REDUCTASE.
CC APPEARS TO ACT AS A METABOLIC STRESS-SENSING PROTEIN KINASE
CC SWITCHING OFF BIOSYNTHETIC PATHWAYS WHEN CELLULAR ATP LEVELS ARE
CC DEPLETED AND WHEN 5'-AMP RISES IN RESPONSE TO FUEL LIMITATION
CC AND/OR HYPOXIA. THIS IS A CATALYTIC SUBUNIT.
CC -!- SUBUNIT: HETEROTRIMER OF A CATALYTIC SUBUNIT, A BETA AND A GAMMA
CC NON-CATALYTIC SUBUNITS.
CC -!- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
CC SNF1 SUBFAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch)
CC -----
CC EMBL; U06454; AAA64745.1; --
CC ENBL; AL035705; CAC17574.1; --
CC PIR; S51025; S51025.
CC HSP; Q63450; 1A06.
CC Genew; HGNC:9377; PRKAA2.
CC MIM; 600497; --
CC GO; GO:0006468; P:protein amino acid phosphorylation; TAS.
CC GO; GO:0007165; P:signal transduction; TAS.
CC InterPro; IPR000719; Prot_kinase.
CC InterPro; IPR002290; Ser_thr_pkinase.
CC Pfam; PF00669; pkinase; 1.
CC ProDom; PD000001; Prot_kinase; 1.
CC SMART; SM00220; S_TKc_1.
CC PROSITE; PS00107; PROTEIN KINASE ATP; 1.
CC PROSITE; PS00108; PROTEIN KINASE ST; 1.
CC PROSITE; PS50011; PROTEIN KINASE DOM; 1.
CC Transferase; Serine/threonine-protein kinase; Fatty acid biosynthesis;
CC Phosphorylation; ATP-binding; Multigene family.
CC DOMAIN 16 268 PROTEIN KINASE.
CC NP_BIND 22 30 ATP (BY SIMILARITY).
CC BINDING 45 45 ATP (BY SIMILARITY).
CC ACT_SITE 139 139 BY SIMILARITY.
CC MOD_RES 172 172 PHOSPHORYLATION (AUTO-) (POTENTIAL).
CC CONFLICT 180 180 A -> T (IN REF. 1).
CC CONFLICT 271 271 D -> G (IN REF. 1).
CC CONFLICT 403 404 HL -> RQ (IN REF. 1).
CC SEQUENCE 552 AA; 62319 MW; C46A6FCD15104975 CRC64;

Query Match 20.7%; Score 725.5; DB 1; Length 552;
Best Local Similarity 29.4%; Pred. No. 7.2e-30;
Matches 194; Conservative 102; Mismatches 221; Indels 143; Gaps 21;

Qy 16 VGPYRLEKTLGKQGLVGLVCHVCTCKVAIKIVNRKLSVLMKVEREAILKLE 74
Db 13 IGHVLDLTGLVGTGFGKVGKIGEHQLTGHKVAIVKILNRQKIRSLDVVGKIKLEQNLKLF 72
Qy 75 HPHVLKLDHVDYENKKYLVLVLEHVSGBELFDYLVKKGRLTPKEARKFRQIISALDFCHS 134
Db 73 HPHIILKLVISTPTDFFWMEVVSGBELFDYICKGRVEEMEARLFFQILSANDYCHR 132
```


GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 26, 2003, 11:57:16 ; Search time 18.3229 Seconds
(without alignments)
3506.035 Million cell updates/sec

Title: US-10-054-579-2

Perfect score: 3497

Sequence: 1 MTSTGKDGGAHQAGYGVYR.....TNCMEMTGRSLKGIIPKS 668

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283308 seqs, 96168682 residues

Total number of hits satisfying chosen parameters: 283308

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

PIR_76:*

1: pir1:*

2: pir2:*

3: pir3:*

4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1743	49.8	887	2 T20941	hypothetical prote
2	795.5	22.7	745	2 G01025	serine/threonine p
3	790	22.6	401	2 B90120	SNF1-related prote
4	790	22.6	1518	2 S37928	probable purine nu
5	785.5	22.5	512	1 J01446	serine/threonine-s
6	772.5	22.1	504	2 T10449	probable serine/th
7	771	22.0	512	2 T52633	serine/threonine-s
8	769.5	22.0	774	2 I48609	probable serine/th
9	764	21.8	576	2 T41587	probable carbon ca
10	763.5	21.8	511	1 A56009	serine/threonine-s
11	763	21.8	562	2 T29858	hypothetical prote
12	759.5	21.7	713	2 S27966	probable serine/th
13	749	21.4	798	2 J07500	qik protein - chic
14	748	21.4	504	2 T07415	probable serine/th
15	736	21.0	1192	2 T18611	probable serine/th
16	735	21.0	1246	2 G98287	protein H39823.1
17	734.5	21.0	512	2 T07788	probable serine/th
18	734	21.0	473	1 S59941	serine/threonine-s
19	734	21.0	775	2 T38929	changed division r
20	729.5	20.9	552	1 A53621	[hydroxymethyl]glut
21	729.5	20.9	915	2 S74283	probable protein k
22	729	20.8	1142	2 S59359	GIN4 protein - yea
23	725.5	20.7	552	1 S51025	[hydroxymethyl]glut
24	722.5	20.7	472	2 B90100	SNF-related kinase
25	720.5	20.6	633	1 A26030	serine/threonine-s
26	718	20.6	602	2 S72513	FOG2 protein - yea
27	704.5	20.1	481	2 I49072	protein kinase - m
28	704.5	20.1	513	1 S60303	serine/threonine-s
29	697	19.9	1398	2 T13741	hypothetical prote

30	688	19.7	502	1 A41361	serine/threonine-s
31	686.5	19.6	513	1 S60304	serine/threonine-s
32	680	19.4	510	2 T04145	serine/threonine p
33	660	18.9	1558	2 T29253	hypothetical prote
34	647	18.5	891	2 T40503	protein kinase kin
35	646.5	18.5	651	2 S52244	p69893 protein - A
36	642.5	18.4	480	2 A86427	probable serine/th
37	637	18.2	891	2 A38903	protein kinase 1 -
38	634	18.1	622	1 S44859	serine/threonine-s
39	630	18.0	726	2 T33998	hypothetical prote
40	624.5	17.9	441	2 C84667	probable protein k
41	622	17.8	440	2 T14736	probable serine/th
42	618	17.7	441	2 E85362	hypothetical prote
43	617	17.6	440	2 T14735	probable serine/th
44	609	17.4	421	2 T48202	protein kinase AK2
45	606	17.3	1064	2 S52687	serine/threonine-s

ALIGNMENTS

RESULT 1

T20941

hypothetical protein F15A2.6 - Caenorhabditis elegans

C:Species: Caenorhabditis elegans

C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 18-Feb-2000

C:Accession: T20941

R:Gregory, J.

submitted to the EMBL Data Library, March 1996

A:Reference number: Z19349

A:Accession: T20941

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-887 <WIL>

A:Cross-references: EMBL:Z70207; PIDN:CAA94127.1; GSPDB:GN00028; CESP:F15A2.6

A:Experimental source: clone F15A2

C:Genetics:

A:Gene: CESP:F15A2.6

A:Map position: X

A:Introns: 32/1; 63/3; 92/2; 139/2; 189/3; 328/2; 448/2; 516/3; 604/2; 684/3; 735/2; 777/

Query Match 49.8%; Score 1743; DB 2; Length 887;

Best Local Similarity 51.1%; Pred. No. 9.6e-61;

Matches 380; Conservative 75; Mismatches 138; Indels 150; Gaps 17;

Qy	13	AQYGVYRLEKTLGKGTGLVKGVCVTCQKVAIKIVNREKLSVLMKVEREIALKL	72
Db	14	AQYCGPYKLEKTLGKGTGLVKTGTHCITGRKVAIKIVNKEKLSVLMKVEREIALMKL	73
Qy	73	IEHPHVLKLDVYENKYLVLVLEHVSGGELFDYLVKGRLLTPKEARKFFRQIISALDFC	132
Db	74	IEHPHVLHLYDVYENKYLVLVLEHVSGGELFDYLVKGRLLTPKEARKFFRQIISALDFC	133
Qy	133	HSHSICHRLDKPENLLDEKNNIRIADFGMASIQVGSLLLETSCGSPHYACPEVIRGEKY	192
Db	134	HAHNIHCHRLDKPENLLDERNNIKVADFGMASIQVGSLLLETSCGSPHYACPEVIRGEKY	193
Qy	193	DGRKADVWSCGVTLFALLVGLPDDNLRQLLEKVKRGVFFHMPHPTPPCOSILRMIE	252
Db	194	DGRKADVWSCGVTLFALLVGLPDDNLRQLLEKVKRGVFFHMPHPTPPCOSILRMIE	253
Qy	253	VDAAARLTLEHIOKHIIWYIG-KGNEPEPEOPIPRKQVIRSLPLEDIDPDVLSMHSLGC	311
Db	254	VDGKRYSLADVFKHPVMSGTTKADPELELPMQOVQVTHVIPGEDSDIDPDVLRHMCNLCG	313
Qy	312	FRDRNKLQDLLSEENQEKVIYFLLLDRKERVPSQDEEDLPFRNEI-----DPPR	362
Db	314	FKQKQLINELLSFKHNTKKNVYFLLLDRKRRRPAQEDD-----TEIVLRGAQNNDDPK	368
Qy	363	KRVDSPLNLR-----HGKRR-----PERKSNEVLVSVD	390
Db	369	KRTDSSRTSRYPMGSIADGSPFINPRKTYGRNQRKSHSLGSGRHSSTPSRSTRDLFGSS	428

391 GGS-----PVPARRAIEAQA----- 406
429 SGSSARAGEDDRGRASASTNHYHTQVDPQTAEARHVRDAQERRESDSGRGS 488
407 ---QGRSRISGASGLSTPLSSPRVTPHPSRGSPPLTPKGTPTVHTPKES---PAGTP 460
489 SRKESKORSKDSKSSSSCKNDASSTSSVPH-----KSPSPVMSSESVVSSSTM 537
461 NPTPPSS-----PSVGVV--PWARLINSIKNSFLGSPRHRRLQVPTPEMSNLT 509
538 NSTNSSTNSLIAGNSQTSIGTSGTSPWRSKLNINIKFLGTPTFRHRK-----MSGT 589
510 PESSPE-----LAKKSWFGNF--ISLEKEQIFVVIKDKPLSSIKADIVHAFLSI 557
590 AESDESUSQMDITDLYKSKSWFGLASMSVERDTHCVPGVGTLSIKAEILRAFLOI 649
558 PLSHSHVISQTSFRAEYK--ATGPAVQKPVKQVNDITYTE-----GGEAQKENGYSV 610
650 HELSHSVVGQNCFRVEYKRGPTVGVSVFSGIKKNNVDIIPSPQQWIAGETP-----TYVV 705
611 TFTLLSGSPRRFRKRVETIQAO 633
706 QFVLLAGVRRFRKRLVEHLSAIL 728

RESULT 2

G01025
serine/threonine protein kinase - human
C;Species: Homo sapiens (man)
C;Date: 21-Dec-1996 #sequence_revision 06-Jun-1997 #text_change 16-Dec-1998
C;Accession: G01025
R;Navarro, E.
submitted to the EMBL Data Library, April 1996
A;Reference number: H00564
A;Accession: G01025
A;Status: preliminary; translated from GB/EMBL/DBJ
A;Molecule type: mRNA
A;Residues: 1-745 <NAV>
A;Cross-references: EMBL:X97630; NID:g1310674
C;Superfamily: unassigned Ser/Thr or Tyr-specific protein kinases; protein kinase homolog
F;18-271/Domain: protein kinase homology <KIN>

Query Match 22.7%; Score 795.5; DB 2; Length 745;
Best Local Similarity 29.1%; Pred. No. 3.4e-24;
Matches 224; Conservative 120; Mismatches 243; Indels 183; Gaps 21;
QY 10 AQHAQYGVYREKTLGKGTGLVKGVCVTCQKVAIKIVNREKLSVLMKVEREIAI 69
DB 11 ADEQPHIGNYRLTKTIGKGNFAKVLARHILTKGEVAVKIIDKTLQNSSSLQKLFREVRI 70
QY 70 LKLIHHPVHLKLDVYENKYLVLVHVSGLGELFDYLVKKGRLTPKARKFFQIISAL 129
DB 71 MKVLNHPNIVLFEVETETKLYLVMEYASGEVFDYLVHGRMKKEARAKFRQIVSAV 130
QY 130 DFCSHSHCHRDLPENLLDEKNIRIADFGMASLQVSDLSLETSCGSPHYACPEVIRG 189
DB 131 QYCHQKFIVRDLKAENLLDADNWKIADFGSNEFTFGNKLTDFCGSPYAPAEPLFQ 190
QY 190 EKYDGRKADVWSCGVILFALLVGLPFDNLRQLLEKVKGVFHMHPFIIPDCQSLIRG 249
DB 191 KKYDGPEDVWMSGLVILTVLSGSLFPDQNLKELRERVLRGKVRIPFYMSTDCENLLK 250
QY 250 MIEVDAARLTLEHIOKHIWYIGGKNEPEPEQIPRKVQIRSLPSLEDDIPDVLDSHSL 309
DB 251 FLILNPSKRGTLQIMKORWNVGHEDD-----LKPYPPLPDYK--DPRTELWMSM 302
QY 310 GCFDRNKLQDLLSEENQERMYIFLLDLDRKERYFSQDEEDLPPRNEID----- 359
DB 303 GYTREE---IQDSLQGVYNEWATYLLGYKSSLEGGDTITLPRPSADLTNSSAQPPS 359
QY 360 -----PPRKVDS-----PMLNRHGK-----RRPE-----RKM----- 383
DB 360 HKVQSVSNPKQRFRSDQAGPAIPTSNYSKTKTSNNAENKPRREDRESGRKASTAKV 419

QY 384 -----EVLVSTDG---GSPVPARRAIEWA---OHGORSRSISG---A 416
DB 420 PASPLPGLERKKTTPPTSTNSVLSTSTNRNSPDLERASLCQASIQNGKDLTWPGSRA 479
QY 417 SSGLSTSPSSPRVTPH-----PSRGSPPLTPKGTPTVHTPKESPA--- 457
DB 480 STASASAGVSAARPRQHOKSMGSHVHNKASGLPPTESNCEVPRPSTAPQRPVVASPSAH 539
QY 458 -----GTNPPT-----PPSPSVGVVFWARLNS 481
DB 540 NISSSGCAPDRTPRGVSSRSTFHAGQLRQVRDQNLPGYVTPASPSGHSQGRGASGS 599
QY 482 IKNSFLGSPRFRHRRLQVP--TPEMSNLT-----ESSPELAKKSW--FG 523
DB 600 IFSKF--TSKFRVRLNLEPESKDRVETLPHVVVSGNDKEKEEFREAKPRSLRFTWSMK 657
QY 524 NFISSLEKEQIFVVIKDKPLSSIKADIVHAFLSIPSLSHSVISQTSFRAEYKATGGPAVF 583
DB 658 TTSSMEPNEMREIRKVLNDANSQSELHEKYMLL--CMHGTPGHEDF----- 702
QY 584 QRPVKQVDITYTEGGEAQKENGISVTFLLSGSPRRFRKRVETIQAO 633
DB 703 ---VQMEVEVC-----KLPRLSLNGVRFRKRSIGTSMFKNITASKIANEL 743

RESULT 3

B90120
SNF1-related protein kinase [imported] - Guillardia theta nucleomorph
C;Species: nucleomorph Guillardia theta
A;Note: a nucleomorph is the vestigial nucleus of a eukaryotic endosymbiont
C;Date: 10-May-2001 #sequence_revision 10-May-2001 #text_change 24-May-2001
C;Accession: B90120
R;Douglas, S.; Zauner, S.; Fraunholz, M.; Beaton, M.; Penny, S.; Deng, L.T.; Wu, X.; Reit
Nature 410, 1091-1096, 2001
A;Title: The highly reduced genome of an enslaved algal nucleus.
A;Reference number: A99082; PMID:11323671; PMID:11323671
A;Accession: B90120
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-401 <DOU>
A;Cross-references: GB:AF083031; NID:g13794278; PIDN:AAK39655.1; GSPDB:GN00152
C;Genetics:
A;Gene: kin(snf2)
A;Map position: 3
A;Genome: nucleomorph
C;Keywords: nucleomorph

Query Match 22.6%; Score 790; DB 2; Length 401;
Best Local Similarity 42.7%; Pred. No. 3.1e-24;
Matches 159; Conservative 87; Mismatches 102; Indels 24; Gaps 7;

QY 14 QYGVYRLEKTLGKGTGLVKGVCVTCQKVAIKIVNREKLSVLMKVEREIAI 71
DB 10 RFVGYLDTLGVGATGKVLAKHSRTKEKVGKVIKKDLFYDKPSLRKIQREISVMK 69
QY 72 LIHHPVHLKLDVYENKYLVLVHVSGLGELFDYLVKKGRLTPKARKFFQIISALDF 131
DB 70 LMFHPVHKIYDVEPSKYLFLIIEVASKGELENYLVEKLENRALKAFHEIISLEY 129
QY 132 CHSHSCHRDLPENLLDEKNIRIADFGMASLQVSDLSLETSCGSPHYACPEVIRGEK 191
DB 130 CHKHICHRLDKLENLLDMKQLKIADFGMASLSIPNIMLTKFCGSPHYASPEVWSNEP 189
QY 192 YDGRKADVWSCGVILFALLVGLPFDNLRQLLEKVKGVFHMHPFIIPDCQSLIRG 249
DB 190 YNGIKADINSGCIIYALVVGKLPYDEENDNMKLFNKFIRFPPRIIHPDCRDLIRA 249
QY 250 MIEVDAARLTLEHIOKHIWYIGGKNEPEPEQIPRKVQIRSLP-----SLEDIPDVL 303
DB 250 LLTINPKRITIDIKKHGPHY---KSPPLPES---CRISQNLSPASVQNPINPDEII 303
QY 304 DSMHSLGCFDRNKLQDLLSEENQERMYIFLLDLDRKERYFSQDEEDLPPRNEIDPPFRK 363

Db 304 SFLLPLQVQDKTKTCGILSTKPSFLVLYRQLEWRKQKMDSVRSNLI---NEANFRK 360
Qy 364 RVDSPMLNRHGK 375
Db 361 K-----LFGQSK 367

RESULT 4
S37928
probable purine nucleotide-binding protein YKL101w - yeast (Saccharomyces cerevisiae)
N:Alternate names: protein YKL433
C:Species: Saccharomyces cerevisiae
C>Date: 03-May-1994 #sequence_revision 03-May-1994 #text_change 24-Sep-1999
C:Accession: S37928; S39084
R:Cherret, G.; Fukuhara, H.; Bolotin-Fukuhara, M.; Daignan-Fornier, B.; Pallier, C.; Puzo
submitted to the Protein Sequence Database, March 1994
A:Reference number: S37920
A:Accession: S37928
A:Molecule type: DNA
A:Residues: 1-1518 <CHR>
A:Cross-references: EMBL:Z28101; NID:G486168; PIDN:CAA81941.1; PID:G486169; MIPS:YKL101w
A:Experimental source: strain S288C
R:Pallier, C.; Valens, M.; Puzos, V.; Fukuhara, H.; Cheret, G.; Sor, F.; Bolotin-Fukuhara
Yeast 9, 1149-1155, 1993
A:Title: DNA sequence analysis of a 17 kb fragment of yeast chromosome XI physically loc
protein kinases.
A:Reference number: S39084; MUID:94078677; PMID:8256524
A:Accession: S39084
A:Status: translation not shown
A:Molecule type: DNA
A:Residues: 1-1518 <PAL>
A:Cross-references: EMBL:X71133; NID:G431205; PIDN:CAA50456.1; PID:G431215
A:Experimental source: strain S288C
C:Genetics:
A:Gene: SGD:HS11
A:Cross-references: SGD:S0001584; MIPS:YKL101w
A:Map position: 11L
C:Superfamily: unassigned Ser/Thr or Tyr-specific protein kinases; protein kinase homolog
C:Keywords: ATP; P-loop; purine nucleotide binding; serine/threonine-specific protein ki
F:79-369/Domain: protein kinase homolog <KIN>
F:79-86/Region: nucleotide-binding motif A (P-loop)
F:87-95/Region: protein kinase ATP-binding motif
F:85/Binding site: ATP/GTP (lys) #status predicted

Query Match 22.6%; Score 790; DB 2; Length 1518;
Best Local Similarity 30.1%; Pred. No. 1.1e-23;
Matches 239; Conservative 107; Mismatches 265; Indels 182; Gaps 21;

Qy 16 VGPYRLKTLGKGTGLVGLGVHCVTCQKVAIKIYNREKL----- 55
Db 78 VGPWKLKTLGKSGSRVRLAKNMTGQLAAIKIVPKKAFVHCSSNGTVPNSYSSMVT 137

Qy 56 -----SESVLMK-----VRETAALKLEHPVHLKHDVYENKKYLYLVLEHS 99
Db 138 SNVSPSTASREHNSHTNPYGIERIEVIMKLSHTNMALFEWENKSELYLVLEYVD 197

Qy 100 GGEFLDYLVKGLRTPKZARKFFROIISALDPCSHSICHRDLKPENLLDEKN-NIRIA 158
Db 198 GGEFLDYLVKGLRTPKZARKFFROIISALDPCSHSICHRDLKPENLLDEKN-NIRIA 257

Qy 159 DFGMASLQVDSLLTSCGSPHYACPEVIRGEKYDGRKADVMSCGVILFALLVGLPFD 218
Db 258 DFGMAELPNKLLTKSCGSPHYASPEIVMGPRYHGGPSDWSGVILFALLTGHLPND 317

Qy 219 DNLRLLEKVKGVFMHPHFPPDQCQLLRGMIEVDAARRLTLEHQHIWYGGKNEPE 278
Db 318 DNIKLLKLVQSGKYQMPNSLSSEARDLISLVIDPEKRITTOELKH--PLIKKYDDL 375

Qy 279 PEQPIPRKVO-----IRSLPLEDIDPDVLDMSHSLGCFDRN 316
Db 376 PVNVLRLKRWKDNMARGNSDLHLNNSVPSIVTLHSGKEIDESILRSLOILMHGVSRE 435

Qy 317 KLLQDLILSEENQKMYLFLLLDRKERY-----PSOEDLPPRNEIDPPKRVDSPLN 371
Db 436 LITAKLLQKPMSEKLFYSLLQYKQHSISLSSSENKKSATSESSVNEPFIASKTAN 495

Qy 372 RHGKRPER--KSMEVLVS--TDGGSVPAPRAIE-----MAHQGRSRS----- 412
Db 496 NTGLRSENNDVKTLSLEIHSIEDTSTVNQNNAITGVNTINAPVLAQKQSIFNTLSQPE 555

Qy 413 -----ISGASSG-----LSTSPLSPRVTPHPSPR 437
Db 556 SDKAEBAVTLPAIPFNAASSRIFNSVTSISRRRRRLRLSNLSLSASTSRSTVH 615

Qy 438 GSPLTPKGTGVHHPKESPAGTPTNTPSPSPSYGVGVPWRARLNSIKNSFLGSPRFHRKL 497
Db 616 DNEMLPQ-LPKSPSYLSRRRAIHASPSKSIH-----KSLSRKNIAATVAARRTL 666

Qy 498 QVTPPEMSNLTPESPPELAKKWFNGFISLEKEQIFVVIKPKPLSSIADIVHAFLSI 557
Db 667 Q-----NSASKRSYLSQSTSKRS-----LNLND-----LLVFDPLPSKPKPASENVKSE 712

Qy 558 PSLSHSVISOTSFA-----EYKATGPAVFOKPVKFOVDITVTE 597
Db 713 P-----HSLESDDFEILCDQLFGNALDRILEEEDNEKERDQORQNDTKSSADTTIS 769

Qy 598 GGEAQRKEN-----GIYSVTFTLLSGPSRRFRV-----VETIQAOQLLSTHDPAA 642
Db 770 GVSTNKENEGPEYPTKIEKNQFNMSYKPSNMGLSFPPIFEKENTLSSSVLEQKPKRA 829

Qy 643 QHLSDDTTCNMEMM 655
Db 830 A-LSDDITNSFNKM 841

RESULT 5
JC1446
serine/threonine-specific protein kinase (EC 2.7.1.-) AK21 - Arabidopsis thaliana
N:Alternate names: protein kinase SNF1 homolog
C:Species: Arabidopsis thaliana (mouse-ear cress)
C>Date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 11-Jun-1999
C:Accession: JC1446; S52266; S66334
R:LeCuen, L.; Thomas, M.; Bianchi, M.; Halford, N.G.; Kreis, M.
Gene 120, 249-254, 1992
A:Title: Structure and expression of a gene from Arabidopsis thaliana encoding a protein
A:Reference number: JC1446; MUID:93013041; PMID:1339373
A:Accession: JC1446
A:Molecule type: DNA
A:Residues: 1-512 <LEG>
A:Cross-references: GB:M93023; NID:G166599; PIDN:AAA32736.1; PID:G166600
R:Thuemmler, F.; Kirchner, M.; Teuber, R.; Ditttrich, P.
submitted to the EMBL Data Library, May 1995
A:Description: Differential accumulation of the transcripts of 22 novel protein kinase ge
A:Reference number: S58256
A:Accession: S58266
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 144-198 <THU>
R:Thuemmler, F.; Kirchner, M.; Teuber, R.; Ditttrich, P.
Plant Mol. Biol. 29, 551-565, 1995
A:Title: Differential accumulation of the transcripts of 22 novel protein kinase genes ir
A:Reference number: S66314; MUID:96123233; PMID:8534852
A:Accession: S66334
A:Molecule type: DNA
A:Residues: 144-198 <TH2>
A:Cross-references: EMBL:X86966; NID:G928909; PIDN:CAA60529.1; PID:G928910
C:Comment: This enzyme plays an important role in a signal transduction cascade regulati
C:Genetics:
A:Gene: AK1n10; AK21
A:Introns: 64/1; 125/3; 186/3; 292/3; 322/3; 350/3; 396/3; 475/3
C:Function:
A:Description: catalyzes the formation of peptidyl-serine-phosphate or peptidyl-threonine
C:Superfamily: AMP-activated protein kinase; protein kinase homology
C:Keywords: ATP; magnesium; phosphotransferase; serine/threonine-specific protein kinase


```

A;Name: 10.004
A;Map position: X
          2      1      0

```

C:Keywords: ATP; phosphotransferase; serine/threonine-specific protein kinase
F:54-307/Domain: protein kinase homology <KIN>
F:62-70/Region: protein kinase ATP-binding motif

Query Match

Best Local Similarity 29.6%, Pred. No. 8e-23;
Matches 213; Conservative 110; Mismatches 246; Indels 151; Gaps 18;

```
QY      10 AQAQYVGPYRLEKTLKGQGTGLVKGVCVTCOKVAIKIWNREKLSGVLMVKYVERETAI 69
      DB      47 ADSQPHIGNYRLKTLKTGKGNFAKVKLARHILTGREVAIKIIDKTLQNPSTSLQKLFREVI 106
QY      70 LKLEIHPHVAKLHDVYENKKYLYLVLEHVSQGGELFDYLVKGRUTLPKEARKFFQIISAL 129
      DB      107 MKTLNHNPIVVKLFEVITQKTLYLIMEYASGGKVFYDLVAHGRMKEKARSKFRQIVSAV 166
QY      130 DFCHSHSICHRDLKPENLLDKNKRIADFGMASLQVGDSLSTSCGSHVACPEVIRG 189
      DB      167 QYCHQKGIIVARDLKAENLLLDADWNIIKIADRGFSNEFTVGSKLDTFCGSPPYAAPELFG 226
QY      190 EKYDGRKADYVSCGVIIFALLVAGALPDDEDDNLRLQLLEKVKRGVFMHPHFIIPDCQSLLRG 249
      DB      227 KKYDGPVEDVMSGLVILVTVSGSLPPDQGNLKELRERVLGKYRIIPPYMTDCENLLKR 286
QY      250 MIEVDAARLLTLEHIOHIIWYIGKNPEPEPPQPIPRKVOIKRSLPSELDIDPDVLDMSHSL 309
      DB      287 FLVNLPIKRGTLQEQIMKDRWINAGHEEDLKPFEV-----PELDISDQKRIDIMVGM 338
QY      310 GCFDRNRKLLQDLLSEENQEKMIYFELLDRKERYPSQDEEDLPPEINIDPPRKRVDSPM 369
      DB      339 GYSGQE-----IQESLSKMKYDITATYLLGRKS-----SEVRP-----SSD 376
QY      370 LNRHGRKRPKRKSMVELSVTDD-----GGSPVPARRAI-----EMAQHGQRS 410
      DB      377 LNNSTGQSPHHKVRQSVSSQKQRRYSHDAGGPGIPSVVAYPKRSQTSSTADSLKEDGIGSS 436
QY      411 RSTSGASSGLSTPLSPSPRVTPHPSPRGSPPLTPKPGTPVHTPKGSPA-----457
      DB      437 RKSTGSAVGGKGIAPASPMGNASPNKADIPERKKSST-VFSSNTASGCMTRRNTYVCS 495
QY      458 -----GTPNPPTPPS-----SPSVGGVPWEARL---NSIKNSFLGSPRF 492
      DB      496 ERTTDDRHSHVIQNGKENSITPDQRTPVASTHISISAATPDRIRPRGTASRSTFHGQPR- 554
QY      493 HRRKLQVPTPEEMSNLTPSSPELAKXS-----WFGNFIS-LEKBEQIFVVVKD-----K 541
      DB      555 ERRTATYNGPPASPSLGSHEATPLQSTRSGSTTLFSLKTSKLTSRNVSAQKQDENKEAK 614
QY      542 PLSSIKADIVHAFILSISLSHSVISQTSFRAEYKATGG--PAVFQPKPVKFOVD-----592
      DB      615 PRS-----LRFWTSMKTTSSMDPGDMWRREIRKVLDDANNCDYE 651
QY      593 -----ITYTGEGAQKEN-----GIYSVTFTLLSGPSRRFRKRVETIQAOQL 633
      DB      652 QRERFLFCVHGDDGHAENLVQWMEVCKLPRLSLNGVTFRKISGTSIAFKNIASKIANEL 711
```

RESULT 13
JC7500
qik protein - chicken
N;Alternate names: Qin-induced kinase
C;Species: Gallus gallus (chicken)
C;Date: 17-Nov-2000 #sequence_revision 17-Nov-2000 #text_change 17-Nov-2000
C;Accession: JC7500
R;Xia, Y.; Zhang, Z.; Kruse, U.; Vogt, P.K.; Li, J.
Biochem. Biophys. Res. Commun. 276, 564-570, 2000
A;Title: The new serine:threonine kinase, Qik, is a target of the qin oncogene
A;Reference number: JC7500
A;Contents: Embryo fibroblasts
A;Accession: JC7500
A;Molecule type: mRNA
A;Residues: 1-798 <XIA>
A;Cross-references: GB:AF219232
C;Comment: This protein, a member of the AMPK/SNF1 family of serine/threonine X
C;Genetics:
A;Gene: qik
C;Keywords: protein kinase

RESULT 14
T07415
probable serine/threonine-specific protein kinase (EC 2.7.1.1.-) PKIN1 - potato
N:Alternate names: SNF1-related protein kinase
C:Species: Solanum tuberosum (potato)
C:Date: 14-May-1999 #sequence_revision 14-May-1999 #text_change 20-Jun-2000
C:Accession: T07415
R:Halford, N.G.
submitted to the EMBL Data Library, March 1996
A:Reference number: Z16022
A:Accession: T07415
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-504 <HAL>
A:Cross-references: EMBL:X95997; PIDN:CAA65244.1
C:Genetics:

Search completed: November 26, 2003, 12:02:38
Job time : 21.3229 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 26, 2003, 11:49:36 ; Search time 36.6897 Seconds
(without alignments)
2630.326 Million cell updates/sec

Title: US-10-054-579-4

Perfect score: 3188

Sequence: 1 MKVEREAILKIEHPHLK.....TNCMEMTGRLSKGIIPKS 608

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1107863 seqs, 158726573 residues

Total number of hits satisfying chosen parameters: 1107863

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A Geneseq 19Jun03.*

1: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1980.DAT.*
2: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1981.DAT.*
3: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1982.DAT.*
4: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1983.DAT.*
5: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1984.DAT.*
6: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1985.DAT.*
7: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1986.DAT.*
8: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1987.DAT.*
9: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1988.DAT.*
10: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1989.DAT.*
11: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1990.DAT.*
12: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1991.DAT.*
13: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1992.DAT.*
14: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1993.DAT.*
15: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1994.DAT.*
16: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1995.DAT.*
17: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1996.DAT.*
18: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1997.DAT.*
19: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1998.DAT.*
20: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1999.DAT.*
21: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA2000.DAT.*
22: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.*
23: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA2002.DAT.*
24: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA2003.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	3188	100.0	608	23	Human kinase NHP #
2	3188	100.0	664	23	Human PKIN-18 prot
3	3188	100.0	668	23	Human kinase NHP #
4	3188	100.0	668	24	Human serine prote
5	3065	96.1	674	23	Human protein kina
6	3028	95.0	585	24	Human serine prote
7	2545.5	79.8	703	24	Human kinase #3.
8	2545.5	79.8	778	24	Human kinase #1.
9	2545.5	79.8	794	22	Human protein kina

10	2545.5	79.8	794	23	AAE16271	Human kinase PKIN-
11	2514.5	78.9	754	24	AAO16504	Human cell cycle-r
12	2386.5	74.9	762	24	ABE98744	Human kinase #2.
13	1754	55.0	851	22	ABE62061	Drosophila melanog
14	1036.5	32.5	301	22	ABE11052	Human HPOK-1 hom
15	766.5	24.0	1349	22	AAE70854	C albicans apoptos
16	695	21.8	514	21	AAE03420	Soybean putative c
17	695	21.8	514	21	AAE04714	Glycine max oil tr
18	693.5	21.8	761	24	ABE96068	Human protein kina
19	689	21.6	768	24	AAE78349	Human protein SEQ
20	688.5	21.6	745	24	AAE33550	Human microtubule
21	688.5	21.6	796	24	ABU11830	Human WDDT polypep
22	686.5	21.5	691	23	AAE19050	Human PAR-1B alpha
23	686.5	21.5	724	23	AAE19051	Human PAR-1B beta
24	686	21.5	504	24	AAE40815	Cucumis sativus oi
25	685.5	21.5	724	24	ABU08527	Human serine/threo
26	685.5	21.5	775	24	AAO16606	Cell cycle-regulat
27	684	21.5	769	23	ABE29666	Human polypeptide
28	682.5	21.4	722	17	AAE98227	Rat neuronal prote
29	681.5	21.4	452	21	AAE36159	Arabidopsis thalia
30	681.5	21.4	454	21	AAE36158	Arabidopsis thalia
31	681.5	21.4	512	21	AAE36157	Arabidopsis thalia
32	681.5	21.4	512	23	AAO17663	A thaliana AKIN11.
33	680	21.3	505	24	ABE40810	Oryza sativa oil t
34	679	21.3	514	24	ABE40816	Glycine max oil tr
35	679	21.3	515	21	AAE03422	Soybean putative c
36	679	21.3	515	24	ABE40716	Glycine max oil tr
37	677	21.2	579	21	AAE03417	Corn putative carb
38	677	21.2	579	24	ABE40709	Zea mays oil trait
39	676.5	21.2	512	23	AAO17664	A thaliana AKIN11
40	675.5	21.2	752	22	AAE11782	Human kinase (PKIN
41	675.5	21.2	752	22	AAE93956	Human polypeptide
42	675.5	21.2	752	23	ABE04433	Human neuronal ser
43	675.5	21.2	752	24	ABE73794	Human MARK protein
44	672.5	21.1	923	22	AAE11773	Human kinase (PKIN
45	672.5	21.1	926	22	AAE65631	Novel protein kina

ALIGNMENTS

RESULT 1

ABE09558
ID ABE09558 standard; Protein; 608 AA.

XX ABE09558;

XX 21-OCT-2002 (first entry)

XX Human kinase NHP #2, SEQ ID NO:4.

XX Human; kinase; NHP; chromosome 11; signal transduction; cytoplasmic;
XX nuclear; brain; pituitary; hypothalamus; adipose; cerebellum;
XX adrenal gland; foetal lung; foetal brain; cancer; drug screening;
XX mental disorder; diagnostic reagent; clinical trial monitoring; cosmetic;
XX nutraceutical; mutation detection; gene expression analysis;
XX chromosome mapping; transgenic animal; cytostatic; gene therapy;
XX enzyme.

XX Homo sapiens.

XX WO200259287-A2.

XX 01-AUG-2002.

XX 22-JAN-2002; 2002WO-US01818.

XX 23-JAN-2001; 2001US-263378P.

XX (LEXI-) LEXICON GENETICS INC.

XX Turner CA, Mathur B;

XX

DR WPI: 2002-599780/64.
DR N-PSDB; ABQ77626.
XX Novel polynucleotide encoding human proteins sharing sequence
PT similarity with animal kinases, useful for drug screening, diagnosis,
PT in gene therapy of disorders and diseases e.g. cancer -
XX
XX
PS Claim 4; Page 39-40; 40pp; English.
XX
CC The invention relates to novel human kinases (AB09557, AAB09558),
CC referred to as NHPs, and to nucleic acids encoding them (ABQ77625,
CC ABQ77626). The NHPs have structural similarity with animal kinases
CC including serine-threonine kinases, carbon catabolite repressing kinases,
CC calcium/calmodulin-dependent protein kinases, and cAMP-dependent protein
CC kinases. The NHPs do not contain consensus signal sequences, indicating
CC that they may be cytoplasmic or nuclear proteins, and are thought to be
CC involved in signal transduction. Polynucleotides encoding NHPs were
CC obtained using human brain, pituitary, hypothalamus, adipose, cerebellum,
CC adrenal gland, foetal lung, and foetal brain cDNAs and primers derived
CC from human genomic DNA. The gene encoding the NHPs is located on
CC chromosome 11. NHP nucleotides and proteins are useful for treating
CC mental, biological or medical disorders including cancer, and for
CC screening compounds useful in the treatment of such conditions. They can
CC also be used as diagnostic reagents, in clinical trial monitoring and in
CC cosmetic and nutraceutical applications. NHP nucleotides can additionally
CC be used in the detection of disease-associated mutations, in the analysis
CC of gene expression, for mapping chromosome 11, for the recombinant
CC expression of NHPs, to generate transgenic animals, in gene therapy,
CC and as part of ribozyme and/or triple helix sequences useful in the
CC modulation of NHP gene expression. The present sequence represents a
CC 608 amino acid NHP (#2).
XX
SQ Sequence 608 AA;

Query Match 100.0%; Score 3188; DB 23; Length 608;
Best Local Similarity 100.0%; Pred. No. 2.2e-228;
Matches 608; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MKVERETAILKLEHPHVILKLDVYENKYLVLVLEHVSQGGELFDYLVKGRLLTPKEARK 60
Db 1 MKVERETAILKLEHPHVILKLDVYENKYLVLVLEHVSQGGELFDYLVKGRLLTPKEARK 60

Qy 61 FFRQIISALDFCHSHSICHRDLKPNLLDEKNNIRIADFGWASIQVSDLSLETSCGSPH 120
Db 61 FFRQIISALDFCHSHSICHRDLKPNLLDEKNNIRIADFGWASIQVSDLSLETSCGSPH 120

Qy 121 YACPEVIRGEKYDGRKADWVSCGVLFPALLVGLPFDNDNLRQLLEKVGKGVFHMHPFIP 180
Db 121 YACPEVIRGEKYDGRKADWVSCGVLFPALLVGLPFDNDNLRQLLEKVGKGVFHMHPFIP 180

Qy 181 PDCOSLLRGMLEVDAAARLTLEHIOKHWYIGGKNEPEPEQIPRKVQIRSLPSLEIDIP 240
Db 181 PDCOSLLRGMLEVDAAARLTLEHIOKHWYIGGKNEPEPEQIPRKVQIRSLPSLEIDIP 240

Qy 241 DVLDSMSLGCFRDRNKLQDLLSEENQEKWYIFLLDRKERYPSQDEDELPNREIDP 300
Db 241 DVLDSMSLGCFRDRNKLQDLLSEENQEKWYIFLLDRKERYPSQDEDELPNREIDP 300

Qy 301 PRKRVDSPLNRHGRKRRPERKSMVLSVTDGSGSPVAPARRAEMAQHQRSRSISGASGL 360
Db 301 PRKRVDSPLNRHGRKRRPERKSMVLSVTDGSGSPVAPARRAEMAQHQRSRSISGASGL 360

Qy 361 STPLSSPRVTPHPSPRGSPPLTPKGTVPHTPKSPAGTPNTPSPSSVGVGVPWARRLN 420
Db 361 STPLSSPRVTPHPSPRGSPPLTPKGTVPHTPKSPAGTPNTPSPSSVGVGVPWARRLN 420

Qy 421 SIKNSFLGSPFRHRRKLVQPTPEEMSNLTPESSPELAKKSWFGNFISLEKEQIFVWIKD 480
Db 421 SIKNSFLGSPFRHRRKLVQPTPEEMSNLTPESSPELAKKSWFGNFISLEKEQIFVWIKD 480

Qy 481 KPLSSIKADIVHAFISPSLSHVSISQTSFRAEYKATGCPAVFQKPVFQVDITYTEGGE 540
Db 481 KPLSSIKADIVHAFISPSLSHVSISQTSFRAEYKATGCPAVFQKPVFQVDITYTEGGE 540

Qy 541 AOKENGIYSVTETLLSGSPSRKRKRVVETIOAQLISTHDPAPAAQHLSDDTNCHEMMTGRLS 600
Db 541 AOKENGIYSVTETLLSGSPSRKRKRVVETIOAQLISTHDPAPAAQHLSDDTNCHEMMTGRLS 600

Qy 601 KCGIIPKS 608
Db 601 KCGIIPKS 608

RESULT 2
AAE21723
ID AAE21723 standard; Protein; 664 AA.
XX
AC AAE21723;
XX
DT 16-JUL-2002 (first entry)
XX
DE Human PKIN-18 protein.
XX
KW Human; kinase; enzyme; PKIN-18 protein; immune system disorder; anaemia;
KW acquired immune deficiency syndrome; thymic hypoplasia; Crohn's disease;
KW asthma; neurological disorder; epilepsy; Charcot-Marie-Tooth disease;
KW AIDS; seizures; cell proliferative disorder; cancer; adenocarcinoma;
KW leukaemia; lymphoma; melanoma; myeloma; sarcoma; developmental disorder;
KW Down's syndrome; gene therapy; protein therapy; cytostatic.
XX
OS Homo sapiens.
XX
FH Key
FT Domain 15..266
FT Domain /note= "Eukaryotic protein kinase domain"
FT Domain 16..257
FT Domain /note= "Protein kinase domain"
FT Domain 17..257
FT Domain /note= "Protein kinase domain"
FT Domain 36..256
FT Domain /note= "Protein kinase domain"
XX
XX WO200218557-A2.
XX
XX 07-MAR-2002.
XX
XX 31-AUG-2001; 2001WO-US27219.
XX
XX 31-AUG-2000; 2000US-229873P.
XX 08-SEP-2000; 2000US-231357P.
XX 14-SEP-2000; 2000US-232654P.
XX 22-SEP-2000; 2000US-234902P.
XX 29-SEP-2000; 2000US-236499P.
XX 06-OCT-2000; 2000US-238389P.
XX 13-OCT-2000; 2000US-240542P.
XX
XX (INCY-) INCYTE GENOMICS INC.
XX
XX Bandnan O, Nguyen DB, Wallia NK, Hafalia AJA, Yao MG, Gandhi AR;
XX Gururajan R, Ding L, Patterson C, Yue H, Baughn MR, Tribouley CM;
XX Thornton M, Elliott VS, Lu Y, Ison CH, Au-Young J, Tang YT;
XX Azimzai Y, Burrill JD, Marcus GA, Zingler KA, Lu DM, Lal PG;
XX Ramkumar J, Warren BA, Kearney L, Pollicky JL, Thangavelu K;
XX Burford N;
XX
XX WPI: 2002-329769/36.
XX N-PSDB; AAD34315.
XX
XX New human kinases, useful for diagnosing, treating or preventing immune
XX system disorders (e.g. Crohn's disease), neurological disorders (e.g.
XX epilepsy), or cell proliferative disorders (e.g. cancers such as
XX leukemia or lymphoma) -
XX
XX Claim 73; Page 181-183; 218pp; English.
XX
XX The present invention relates to human kinases (PKIN) and polynucleotides
CC

CC encoding such proteins. PKIN sequences of the invention are useful for
CC diagnosing, treating or preventing disorders associated with aberrant
CC expression of PKIN, particularly immune system disorders (e.g. acquired
CC immune deficiency syndrome (AIDS), thymic hypoplasia, Crohn's disease,
CC anaemia, asthma), neurological disorders (e.g. epilepsy, Charcot-Marie-
CC Tooth disease or seizures), cell proliferative disorders (e.g. cancers
CC such as adenocarcinoma, leukaemia, lymphoma, melanoma, sarcoma),
CC and developmental disorders (e.g. Down's syndrome). They are also used
CC in gene therapy and protein therapy. The present sequence is human
CC PKIN-18 protein.
XX
SQ Sequence 664 AA;

Query Match 100.0%; Score 3188; DB 23; Length 664;
Best Local Similarity 100.0%; Pred. No. 2.4e-228;
Matches 608; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 MKVEREIAILKLIIEHPHVLKLDVYENKYLVLVLEHVS GGELFDYLVKGR LTPKEARK 60
Db 57 MKVEREIAILKLIIEHPHVLKLDVYENKYLVLVLEHVS GGELFDYLVKGR LTPKEARK 116
Qy 61 FFRQIISALDFCHSHSICHRLDKPENLLDEKNIRIADFGWASIQVGS LLETSCGSPH 120
Db 117 FFRQIISALDFCHSHSICHRLDKPENLLDEKNIRIADFGWASIQVGS LLETSCGSPH 176
Qy 121 YACPEVIRGEKYDGRKADVWSCGVILFALLVGLPFDDNNLRQLLEKVRGVFHPHFIP 180
Db 177 YACPEVIRGEKYDGRKADVWSCGVILFALLVGLPFDDNNLRQLLEKVRGVFHPHFIP 236
Qy 181 PDCQSLLRGMIEVDAAARLTLLEHIQKHIWYIGKNEPEPEQIPRKVQIRSLPSLEDIDP 240
Db 237 PDCQSLLRGMIEVDAAARLTLLEHIQKHIWYIGKNEPEPEQIPRKVQIRSLPSLEDIDP 296
Qy 241 DVLDSMHSIGCFDRNKLQDLSEENQEKMIYFLLDRKERYPSQDEDLPPRNEIDP 300
Db 297 DVLDSMHSIGCFDRNKLQDLSEENQEKMIYFLLDRKERYPSQDEDLPPRNEIDP 356
Qy 301 PRKRVDSPLNLRHGRPRKSRMEVSLVTDGSPVPARRAIEAQMAGORSRSISGASSGL 360
Db 357 PRKRVDSPLNLRHGRPRKSRMEVSLVTDGSPVPARRAIEAQMAGORSRSISGASSGL 416
Qy 361 STSPSSSRVTPHPSPRSGPLPTPKGTVPHTPKESPAGTPNTPPSSPSVGVVWRARLN 420
Db 417 STSPSSSRVTPHPSPRSGPLPTPKGTVPHTPKESPAGTPNTPPSSPSVGVVWRARLN 476
Qy 421 STKNSFLGSPRHRKQLQVTPPEMSNLTPSSPELAKSMFGNFISEKEQIFVWLKD 480
Db 477 STKNSFLGSPRHRKQLQVTPPEMSNLTPSSPELAKSMFGNFISEKEQIFVWLKD 536
Qy 481 KPLSSIKADIVHAFSLISPSLSHSVISQTSFRAEYKATGPAVFPKPVQVDITYEGGE 540
Db 537 KPLSSIKADIVHAFSLISPSLSHSVISQTSFRAEYKATGPAVFPKPVQVDITYEGGE 596
Qy 541 AQKENGIVSVTPTLLSGSRRRPRVETIQALLSTHDPAAQHLSDTNCMEMMTGRLS 600
Db 597 AQKENGIVSVTPTLLSGSRRRPRVETIQALLSTHDPAAQHLSDTNCMEMMTGRLS 656
Qy 601 KCGIIPKS 608
Db 657 KCGIIPKS 664

RESULT 3
ABB09557
ID ABB09557 standard; Protein; 668 AA.
XX
AC ABB09557;
XX
DT 21-OCT-2002 (first entry)
XX
DE Human kinase NHP #1, SEQ ID NO:2.
XX
KW Human; kinase; NHP; chromosome 11; signal transduction; cytoplasmic;

KW nuclear; brain; pituitary; hypothalamus; adipose; cerebellum;
KW adrenal gland; foetal lung; foetal brain; cancer; drug screening;
KW mental disorder; diagnostic reagent; clinical trial monitoring; cosmetic;
KW nutraceutical; mutation detection; gene expression analysis;
KW chromosome mapping; transgenic animal; cycostatic; gene therapy;
XX Homo sapiens.
XX W0200259287-A2.
XX 01-AUG-2002.
XX 22-JAN-2002; 2002WO-US01818.
XX 23-JAN-2001; 2001US-263378P.
XX (LEXI-) LEXICON GENETICS INC.
XX Turner CA, Mathur B;
XX WPI; 2002-599780/64.
XX N-PSDB; ABQ77625.
XX Novel polynucleotide encoding human proteins sharing sequence
XX similarity with animal kinases, useful for drug screening, diagnosis,
XX in gene therapy of disorders and diseases e.g. cancer -
XX Claim 2; Page 37-38; 40pp; English.
XX The invention relates to novel human kinases (ABB09557, AAB09558),
XX referred to as NHPs, and to nucleic acids encoding them (ABQ77625,
XX ABQ77626). The NHPs have structural similarity with animal kinases
XX including serine-threonine kinases, carbon catabolite repressing kinases,
XX calcium/calmodulin-dependent protein kinases, and cAMP-dependent protein
XX kinases. The NHPs do not contain consensus signal sequences, indicating
XX that they may be cytoplasmic or nuclear proteins, and are thought to be
XX involved in signal transduction. Polynucleotides encoding NHPs were
XX obtained using human brain, pituitary, hypothalamus, adipose, cerebellum,
XX adrenal gland, foetal lung, and foetal brain cDNAs and primers derived
XX from human genomic DNA. The gene encoding the NHPs is located on
XX chromosome 11. NHP nucleotides and proteins are useful for treating
XX mental, biological or medical disorders including cancer, and for
XX screening compounds useful in the treatment of such conditions. They can
XX also be used as diagnostic reagents, in clinical trial monitoring and in
XX cosmetic and nutraceutical applications. NHP nucleotides can additionally
XX be used in the detection of disease-associated mutations, in the analysis
XX of gene expression, for mapping chromosome 11, for the recombinant
XX expression of NHPs, to generate transgenic animals, in gene therapy,
XX and as part of ribozyme and/or triple helix sequences useful in the
XX modulation of NHP gene expression. The present sequence represents a
XX 668 amino acid NHP (#1).

Query Match 100.0%; Score 3188; DB 23; Length 668;
Best Local Similarity 100.0%; Pred. No. 2.5e-228;
Matches 608; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 MKVEREIAILKLIIEHPHVLKLDVYENKYLVLVLEHVS GGELFDYLVKGR LTPKEARK 60
Db 61 MKVEREIAILKLIIEHPHVLKLDVYENKYLVLVLEHVS GGELFDYLVKGR LTPKEARK 120
Qy 61 FFRQIISALDFCHSHSICHRLDKPENLLDEKNIRIADFGWASIQVGS LLETSCGSPH 120
Db 121 FFRQIISALDFCHSHSICHRLDKPENLLDEKNIRIADFGWASIQVGS LLETSCGSPH 180
Qy 121 YACPEVIRGEKYDGRKADVWSCGVILFALLVGLPFDDNNLRQLLEKVRGVFHPHFIP 180
Db 181 YACPEVIRGEKYDGRKADVWSCGVILFALLVGLPFDDNNLRQLLEKVRGVFHPHFIP 240
Qy 181 PDCQSLLRGMIEVDAAARLTLLEHIQKHIWYIGKNEPEPEQIPRKVQIRSLPSLEDIDP 240
XX

Db 241 PDCQSLRGMIETVDAARRLTLEHIQKHIWYIGGKNEPEPEQIPRKVQIRSLPSLEDIDP 300
Qy 241 DVLDSMHSGLGCFDRNKLQDLSEENQKMIYFLLDRKERYPSQDEDLPRNEIDP 300
Db 301 DVLDSMHSGLGCFDRNKLQDLSEENQKMIYFLLDRKERYPSQDEDLPRNEIDP 360
Qy 301 PRKRVDSPLNHRHGKRRPERKSMVLSVTDGSPVPARRAIEMAHQGQRSISGASSGL 360
Db 361 PRKRVDSPLNHRHGKRRPERKSMVLSVTDGSPVPARRAIEMAHQGQRSISGASSGL 420
Qy 361 STSPLSSPRVTPHPSGRGSPPLTPKGTVPVHTPKESPACTNPPTPPSPSVGGVFWARLN 420
Db 421 STSPLSSPRVTPHPSGRGSPPLTPKGTVPVHTPKESPACTNPPTPPSPSVGGVFWARLN 480
Qy 421 SIKNSFLGSPFRHRRKLQVPTPEMSNLTPESSPELAKSWFGNFISLEKEBOIFVVIKD 480
Db 481 SIKNSFLGSPFRHRRKLQVPTPEMSNLTPESSPELAKSWFGNFISLEKEBOIFVVIKD 540
Qy 481 KPLSSIKADIVHAFSLSPSLSHSVISQTSFRAEYKATGGPAVFQKPVKQVQDITYTEGGE 540
Db 541 KPLSSIKADIVHAFSLSPSLSHSVISQTSFRAEYKATGGPAVFQKPVKQVQDITYTEGGE 600
Qy 541 AQKNGIYSVTFTLLSGPSRRFRKRVVETIQAOQLLSTHDPAAQHLSDTTNCMEMMTGRLS 600
Db 601 AQKNGIYSVTFTLLSGPSRRFRKRVVETIQAOQLLSTHDPAAQHLSDTTNCMEMMTGRLS 660
Qy 601 KCGIIPKS 608
Db 661 KCGIIPKS 668

RESULT 4

ABG72382
ID ABG72382 standard; Protein; 668 AA.
XX ABG72382;
AC XX
DT 10-FEB-2003 (first entry)
XX
DE Human serine protein kinase KSE336-1.
XX
KW Human; enzyme; chromosome 11p15.5-pter; astrocytoma; meningioma;
KW pancreatic adenocarcinoma; insulin-dependent diabetes mellitus 2;
KW helioid peripapillary chorioretinal degeneration; brain; pancreas;
KW Beckwith-Wiedemann syndrome; congenital hyperinsulinism; KSE336.
XX
OS Homo sapiens.
XX
FN US6455292-B1.
XX
PD 24-SEP-2002.
XX
PF 16-AUG-2001; 2001US-0930181.
XX
PR 16-AUG-2001; 2001US-0930181.
XX
PA (ORIG-) ORIGENE TECHNOLOGIES INC.
XX
PI Shu Y, Fan W, Kovacs KF, Zidanic M, Jay G;
XX
PI WPI; 2003-066233/06.
DR N-PSDB; ABX13583.
XX
XX New isolated polynucleotide coding without interruption for a human
PT KSE336 polypeptide useful for preventing or treating
PT disease/conditions relating to brain and pancreas, e.g. meningioma,
PT insulin-dependent diabetes mellitus 2
XX
XX Claim 1; Fig 2; 3app; English.
PS
XX The invention relates to an isolated polynucleotide (its complement
CC or a sequence 99% similar to it) coding without interruption for a human
CC KSE336 polypeptide, a serine protein kinase, comprising the KSE336-1 and

CC KSE336-2 splice variants appearing as ABG72382 and ABG72383. Also
CC included is a method of identifying an agent that modulates the
CC expression of KSE336 in brain, pancreas, brain progenitor or pancreas
CC progenitor cells comprising: (a) contacting a cell population comprising
CC the cells with a test agent under conditions effective for the test agent
CC to modulate the expression of KSE336; and (b) determining if the test
CC agent modulates the expression of KSE336. The polynucleotides are useful
CC as molecular targets or drug targets, and for detecting, diagnosing,
CC staging, monitoring, prognosticating, preventing or treating diseases or
CC conditions relating to brain and pancreas, such as astrocytoma,
CC meningioma, pancreatic adenocarcinoma, insulin-dependent diabetes
CC mellitus 2, helioid peripapillary chorioretinal degeneration, Beckwith-
CC Wiedemann syndrome or congenital hyperinsulinism. The method and
CC polynucleotides are useful in research, diagnosis, drug discovery,
CC therapy, clinical medicine, forensic science and pathology.
CC The gene for KSE336 is located on chromosome 11p15.5-pter.
CC The present sequence represents the KSE336 variant KSE366-1.
XX

SQ Sequence 668 AA;

Query Match 100.0%; Score 3188; DB 24; Length 668;
Best Local Similarity 100.0%; Pred. No. 2.5e-228;
Matches 608; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MKVEREIALKUIEHPHVLKLDHVVYENKYLVLVLEHVSGBELFDYLVKGRITPKEARK 60
Db 61 MKVEREIALKUIEHPHVLKLDHVVYENKYLVLVLEHVSGBELFDYLVKGRITPKEARK 120
Qy 61 FFRQIISALDFCHSHSI CHRDLPENLLDEKNNIRIADFGWASLOVGSLLTSCGSPH 120
Db 121 FFRQIISALDFCHSHSI CHRDLPENLLDEKNNIRIADFGWASLOVGSLLTSCGSPH 180
Qy 121 YACPEVIRGEKVDGRKADVWSCGVILFALLVGALEPDNDLRLQLEKVRGVFHFHFIP 180
Db 181 YACPEVIRGEKVDGRKADVWSCGVILFALLVGALEPDNDLRLQLEKVRGVFHFHFIP 240
Qy 181 PDCQSLRGMIETVDAARRLTLEHIQKHIWYIGGKNEPEPEQIPRKVQIRSLPSLEDIDP 240
Db 241 PDCQSLRGMIETVDAARRLTLEHIQKHIWYIGGKNEPEPEQIPRKVQIRSLPSLEDIDP 300
Qy 241 DVLDSMHSGLGCFDRNKLQDLSEENQKMIYFLLDRKERYPSQDEDLPRNEIDP 300
Db 301 DVLDSMHSGLGCFDRNKLQDLSEENQKMIYFLLDRKERYPSQDEDLPRNEIDP 360
Qy 301 PRKRVDSPLNHRHGKRRPERKSMVLSVTDGSPVPARRAIEMAHQGQRSISGASSGL 360
Db 361 PRKRVDSPLNHRHGKRRPERKSMVLSVTDGSPVPARRAIEMAHQGQRSISGASSGL 420
Qy 361 STSPLSSPRVTPHPSGRGSPPLTPKGTVPVHTPKESPACTNPPTPPSPSVGGVFWARLN 420
Db 421 STSPLSSPRVTPHPSGRGSPPLTPKGTVPVHTPKESPACTNPPTPPSPSVGGVFWARLN 480
Qy 421 SIKNSFLGSPFRHRRKLQVPTPEMSNLTPESSPELAKSWFGNFISLEKEBOIFVVIKD 480
Db 481 SIKNSFLGSPFRHRRKLQVPTPEMSNLTPESSPELAKSWFGNFISLEKEBOIFVVIKD 540
Qy 481 KPLSSIKADIVHAFSLSPSLSHSVISQTSFRAEYKATGGPAVFQKPVKQVQDITYTEGGE 540
Db 541 KPLSSIKADIVHAFSLSPSLSHSVISQTSFRAEYKATGGPAVFQKPVKQVQDITYTEGGE 600
Qy 541 AQKNGIYSVTFTLLSGPSRRFRKRVVETIQAOQLLSTHDPAAQHLSDTTNCMEMMTGRLS 600
Db 601 AQKNGIYSVTFTLLSGPSRRFRKRVVETIQAOQLLSTHDPAAQHLSDTTNCMEMMTGRLS 660
Qy 601 KCGIIPKS 608
Db 661 KCGIIPKS 668

RESULT 5

AAM47830
ID AAM47830 standard; Protein; 674 AA.
XX

AC ANM47830;
XX 19-FEB-2002 (first entry)
XX Human protein kinase 2246 SEQ ID NO 2.
XX Human; protein kinase 2246; cytostatic; immunomodulator; carcinoma;
KW anti-inflammatory; analgesic; cardiovascular; cancer; sarcoma;
KW cellular proliferation disorder; cellular differentiation disorder;
KW metastatic; haematopoietic disorder; leukaemia; immune disorder;
KW inflammatory disorder; arthritis; autoimmune disease; diabetes mellitus;
KW psoriasis; Crohn's disease; cardiovascular disease; virus; pain;
KW gene therapy.
XX Homo sapiens.
XX WO200181598-A2.
XX 01-NOV-2001.
XX 25-APR-2001; 2001WO-US13784.
XX 25-APR-2000; 2000US-199391P.
XX (MILL-) MILLENNIUM PHARM INC.
XX Meyers R;
XX WPI; 2002-045281/06.
XX N-PSDB; ABA02994; ABA02995.
XX New protein kinase nucleic acid and polypeptide molecules, designated
XX 2246, useful for diagnosing, preventing or treating cancer or a
XX cellular proliferation/differentiation disorders, e.g. carcinoma,
XX sarcoma or leukaemias -
XX Claim 4; Fig 1; 11pp; English.
XX The invention relates to the human protein kinase 2246 gene and the
XX the isolated encoded polypeptide with cytostatic, immunomodulator,
XX anti-inflammatory, analgesic and cardiovascular activity. The 2246
XX nucleic acid and polypeptide are useful for diagnosing, preventing or
XX treating a subject having cancer or a cellular proliferation and/or
XX differentiation disorder or at risk of developing cancer or a cellular
XX proliferation and/or differentiation disorder. In particular, the
XX disorder includes carcinoma, sarcoma, metastatic or haematopoietic
XX disorders (e.g. leukaemias) or cancers of the lung, breast, thyroid, head
XX neck, prostate or genito-urinary tract. The 2246 nucleic acid and
XX polypeptide are also useful for treating immune disorders, e.g.
XX inflammatory (e.g. respiratory inflammation or arthritis), autoimmune
XX disease (e.g. diabetes mellitus, psoriasis, Wegener's granulomatosis,
XX Crohn's disease or Grave's disease), for treating cardiovascular
XX diseases, endothelial cell disorder, viral diseases or pain. The nucleic
XX acid and polypeptide are also useful for evaluating the efficacy of a
XX treatment of cancer or a cellular proliferation and/or differentiation
XX disorder. The nucleic acid is useful for gene therapy. The present
XX sequence is that of the 2246 protein kinase.
XX Sequence 674 AA;
XX
Query Match 96.1%; Score 3065; DB 23; Length 674;
Best Local Similarity 99.7%; Pred. No. 3.5e-219;
Matches 585; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
Qy 1 MKVEREAILKLEPHVYLKLDVYENKYLVLVLEHVSQGELEFDYLVKGRLLTPKARK 60
Db 61 MKVEREAILKLEPHVYLKLDVYENKYLVLVLEHVSQGELEFDYLVKGRLLTPKARK 120
Qy 61 FFRQIISALDFCHSHSICHRDLKPNLLDKNNIRIADFGWASIQVDSILETSCGSPH 120
Db 121 FFRQIISALDFCHSHSICHRDLKPNLLDKNNIRIADFGWASIQVDSILETSCGSPH 180
Qy 121 YACPEVIRGEKYDGRKADVWSCGVILFALLVGPDDNLRQLLEKVKRGVFMHFP 180

Db 181 YACPEVIRGEKYDGRKADVWSCGVILFALLVGPDDNLRQLLEKVKRGVFMHFP 240
Qy 181 PDCQSLLRGMIEVDAAARLLTLEHIQKHIIWYIGGKNEPEPEPIPRKQVIRSLPSLEDIDP 240
Db 241 PDCQSLLRGMIEVDAAARLLTLEHIQKHIIWYIGGKNEPEPEPIPRKQVIRSLPSLEDIDP 300
Qy 241 DVLDSDHSLGCFDRNKLQDLSEENQEKMIYFLLDKERYPSQEDDLPRNEIDP 300
Db 301 DVLDSDHSLGCFDRNKLQDLSEENQEKMIYFLLDKERYPSQEDDLPRNEIDP 360
Qy 301 PRKRVDSPLNLRHGRKRRPERKSMVLSVTDGSGSPVARRAIEAQAHQGRSISGASSGL 360
Db 361 PRKRVDSPLNLRHGRKRRPERKSMVLSVTDGSGSPVARRAIEAQAHQGRSISGASSGL 420
Qy 361 STSPLSSPRVTPHPSPRGSPPLTPKTPVHTPKESPACTNPPTPPSPSGVGVFWARLN 420
Db 421 STSPLSSPRVTPHPSPRGSPPLTPKTPVHTPKESPACTNPPTPPSPSGVGVFWARLN 480
Qy 421 SIKNSFLGSPRRHRRKLOVPTPEMSNLTPESPPELAKKSWFCNFIISLEKEEIQFVVIKD 480
Db 481 SIKNSFLGSPRRHRRKLOVPTPEMSNLTPESPPELAKKSWFCNFIISLEKEEIQFVVIKD 540
Qy 481 KPLSSIKADIVHAFLSIPSLSHSVISQTSFRAFYKATGPAVFPKPKFQVDITYTEGGE 540
Db 541 KPLSSIKADIVHAFLSIPSLSHSVISQTSFRAFYKATGPAVFPKPKFQVDITYTEGGE 600
Qy 541 AQKENGISYVTFLLSGPSRRPKRVVETIQAQLLSTHDPAAQHLSD 587
Db 601 AQKENGISYVTFLLSGPSRRPKRVVETIQAQLLSTHDPAAQHLSE 647
RESULT 6
ABG72383
ID ABG72383 standard; Protein; 585 AA.
XX AC ABG72383;
XX DT 10-FEB-2003 (first entry)
XX DE Human serine protein kinase KSE336-2.
XX KW Human; enzyme; chromosome 11p15.5-pter; astrocytoma; meningioma;
KW pancreatic adenocarcinoma; insulin-dependent diabetes mellitus 2;
KW helicoid peripapillary chorioretinal degeneration; brain; pancreas;
KW Beckwith-Wiedemann syndrome; congenital hyperinsulinism; KSE336.
XX Homo sapiens.
XX US6455292-B1.
XX PD 24-SEP-2002.
XX PF 16-AUG-2001; 2001US-0930181.
XX PR 16-AUG-2001; 2001US-0930181.
XX PA (ORIG-) ORIGENE TECHNOLOGIES INC.
XX Shu Y, Fan W, Kovacs KF, Zidanic M, Jay G;
XX WPI; 2003-066233/06.
XX DR N-PSDB; ABX13584.
XX New isolated polynucleotide coding without interruption for a human
PT KSE336 polypeptide useful for preventing or treating
PT diseases/conditions relating to brain and pancreas, e.g. meningioma,
PT insulin-dependent diabetes mellitus 2 -
XX Claim 1; Fig 2; 34pp; English.
XX The invention relates to an isolated polynucleotide (its complement
CC or a sequence 99% similar to it) coding without interruption for a human
CC

CC KSE336 polypeptide, a serine protein kinase, comprising the KSE336-1 and
CC KSE336-2 splice variants appearing as ABG72382 and ABG72383. Also
CC included is a method of identifying an agent that modulates the
CC expression of KSE336 in brain, pancreas, brain progenitor or pancreas
CC progenitor cells comprising: (a) contacting a cell population comprising
CC the cells with a test agent under conditions effective for the test agent
CC to modulate the expression of KSE336; and (b) determining if the test
CC agent modulates the expression of KSE336. The polynucleotides are useful
CC as molecular targets or drug targets, and for detecting, diagnosing,
CC staging, monitoring, prognosticating, preventing or treating diseases or
CC conditions relating to brain and pancreas, such as astrocytoma,
CC meningioma, pancreatic adenocarcinoma, insulin-dependent diabetes
CC mellitus 2, helicoid papillary choroidretinal degeneration, Beckwith-
CC Wiedemann syndrome or congenital hyperinsulinism. The method and
CC polynucleotides are useful in research, diagnosis, drug discovery,
CC therapy, clinical medicine, forensic science and pathology.
CC The gene for KSE336 is located on chromosome 1p15.5-pter.
CC The present sequence represents the KSE366 variant KSE366-2.
XX
XX
XX Sequence 585 AA;

Query Match 95.0%; Score 3028; DB 24; Length 585;
Best Local Similarity 100.0%; Pred. No. 1.6e-216;
Matches 577; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 32 YLVLEHVSGLGELFDYLVKGRITPKARKEFFROIISALDFCHSHSICHRDLKPNLLDE 91
DB 9 YLVLEHVSGLGELFDYLVKGRITPKARKEFFROIISALDFCHSHSICHRDLKPNLLDE 68
QY 92 KNNIRIADFGMSLQVDSLETSCGSPHYACPEVIRGEKYDGRKADWSCGVILFALLV 151
DB 69 KNNIRIADFGMSLQVDSLETSCGSPHYACPEVIRGEKYDGRKADWSCGVILFALLV 128
QY 152 GALLPDDNNLRQLLEKVKRGVFMHPHPIPDCCSLRGMIIEVDAAARLTLEHIQKIWIYI 211
DB 129 GALLPDDNNLRQLLEKVKRGVFMHPHPIPDCCSLRGMIIEVDAAARLTLEHIQKIWIYI 188
QY 212 GKNPEPEPIPRKQVIRSLPSLEDDPDVLDMSHSLGCFDRDNKLLQDLLSEENQEK 271
DB 189 GKNPEPEPIPRKQVIRSLPSLEDDPDVLDMSHSLGCFDRDNKLLQDLLSEENQEK 248
QY 272 MIYFLLLDKERYPSQDEDLPPRNEIDPPRKRVDSPMLNRHGKRRPERKSMVLSVTDG 331
DB 249 MIYFLLLDKERYPSQDEDLPPRNEIDPPRKRVDSPMLNRHGKRRPERKSMVLSVTDG 308
QY 332 GSPVPARRAEMAQHGORSISGASSGLSTSPSSPRVTPHSPRGSPLTPKGTVPHT 391
DB 309 GSPVPARRAEMAQHGORSISGASSGLSTSPSSPRVTPHSPRGSPLTPKGTVPHT 368
QY 392 PKESPAGTNPPTPPSPSVGGVPWRRLNSIKNSFLGSPRFRHKLQVPTPEMSNLTP 451
DB 369 PKESPAGTNPPTPPSPSVGGVPWRRLNSIKNSFLGSPRFRHKLQVPTPEMSNLTP 428
QY 452 SSPELAKSWFGNFIISLEKEQIFVWIKDKPLSIKADIVHAFISIPSLSHSVISQTSFR 511
DB 429 SSPELAKSWFGNFIISLEKEQIFVWIKDKPLSIKADIVHAFISIPSLSHSVISQTSFR 488
QY 512 ABEYKATGGPVPKPVKQFVDITVTEGEAQKNGIYSVFTLLSGPSRRFRVETIOA 571
DB 489 ABEYKATGGPVPKPVKQFVDITVTEGEAQKNGIYSVFTLLSGPSRRFRVETIOA 548
QY 572 QLLSTHPPAAQHLSDTTNCMMMTGRLSKCGIIPKS 608
DB 549 QLLSTHPPAAQHLSDTTNCMMMTGRLSKCGIIPKS 585

RESULT 7
ABB98745
ID ABB98745 standard; Protein; 703 AA.
XX
XX
AC ABB98745;
XX
DT 20-JAN-2003 (first entry)

XX Human kinase #3.
DE
XX Human; kinase; chromosome 19.
XX
XX Homo sapiens.
OS
XX WO200281670-A1.
PN
XX 17-OCT-2002.
PD
XX 04-APR-2002; 2002WO-US10786.
PF
XX 06-APR-2001; 2001US-282036P.
PR
XX (LEXI-) LEXICON GENETICS INC.
PA
XX Turner CA, Mathur B, Friddle CJ;
PI
XX WPI; 2003-058538/05.
XX N-PSDB; ABV74559.
DR
XX New human kinase proteins useful for diagnosis, drug screening, and
PT clinical trial monitoring, treatment of disorders and diseases, and
PT cosmetic and nutritional applications
PS
XX Claim 5; Page 44-46; 47pp; English.
XX
XX The present sequence is a novel human kinase. The genomic locus encoding
CC the kinase is thought to be on human chromosome 19. The kinase and its
CC coding sequence are useful for diagnosis, drug screening, clinical trial
CC monitoring, treatment of disorders and diseases, and cosmetic and
CC nutritional applications.
XX
XX Sequence 703 AA;
SQ
Query Match 79.8%; Score 2545.5; DB 24; Length 703;
Best Local Similarity 75.6%; Pred. No. 1.5e-180;
Matches 495; Conservative 51; Mismatches 44; Indels 65; Gaps 7;
QY 1 MKVEREIALTKLIEHPHVLKHDVYENKKYLVLEHVSGLGELFDYLVKGRITPKAR 60
DB 1 MKVEREIALTKLIEHPHVLKHDVYENKKYLVLEHVSGLGELFDYLVKGRITPKAR 60
QY 61 FFRQIISALDFCHSHSICHRDLKPNLLDEKNNIRIADFGMSLQVDSLETSCGSPH 120
DB 61 FFRQIISALDFCHSHSICHRDLKPNLLDEKNNIRIADFGMSLQVDSLETSCGSPH 120
QY 121 YACPEVIRGEKYDGRKADWSCGVILFALLVGLALPDDNNLRQLLEKVKRGVFMHPHPI 180
DB 121 YACPEVIRGEKYDGRKADWSCGVILFALLVGLALPDDNNLRQLLEKVKRGVFMHPHPI 180
QY 181 PDCQSLRGMIIEVDAAARLTLEHIQKIWIYIGKNEPEP-EQIP-RKQVIRSLPSLEDI 238
DB 181 PDCQSLRGMIIEVDAAARLTLEHIQKIWIYIGKNEPEP-EQIP-RKQVIRSLPSLEDI 238
QY 239 DPVLDMSHSLGCFDRDNKLLQDLLSEENQEKMIYFLLLDKERYPSQDEDLPPRNEI 298
DB 241 DPVLDMSHSLGCFDRDNKLLQDLLSEENQEKMIYFLLLDKERYPSQDEDLPPRNDV 300
QY 299 DPPRKRVDSFMLNRHGKRRPERKSMVLSVTDG---CGSPVPARRAEMAQHGORSISG 355
DB 301 DPPRKRVDSFMLNRHGKRRPERKSMVLSITDAGGGGSPVPTRRALLEMAQHSQRSVSG 360
QY 356 ASSGLSTSPSSPR-----VTPHP 374
DB 361 ASTGLSSSPSSPRSPVFSFSPFGAGDEARGGSGTSKTQTLPSRGRGGGAGEQPPPP 420
QY 375 SPRGSPLETPKG-----TPVHTPKESPAGTNPPTPPSSP--SVGVPWRRLNS 421
DB 421 SARSTPLPGPGSPSSGGTFLHSLPHTPRASPTGTPTTPPSGGVCGAAWRRLNS 480
QY 422 IKNSFLGSPRFRHKLQVPTPEMSNLTPSSPELAKSWFGNFIISLEKEQIFVVIKDK 481

```
Db 481 INNSFLGSPRRRRKMQVPTAEEMSLTPESPSELAKRSWFGNFIISLDKEQIFLVLDK 540
Qy 482 PLSSIKADIVHAFSLIPSLSHSVISQTSFRAEYKATGPAVFQKPVKQVDITYEGGEA 541
Db 541 PLSSIKADIVHAFSLIPSLSHSVISQTSFRAEYKASGGPSVFQKPVQVDISSSEGEPEP 600
Qy 542 QKE-----NGIYSVTFTLLSGPSRRFRKRVVETIQALLSTHDPAAQHLSDTTN 590
Db 601 SPRDGGGGGIYSVTFTLLSGPSRRFRKRVVETIQALLSTHDPQSVQALADEKN 655

RESULT 8
ID ABB98743 standard; Protein; 778 AA.
AC ABB98743;
XX
DT 20-JAN-2003 (first entry)
DE Human kinase #1.
KW Human; kinase; chromosome 19.
OS Homo sapiens.
XX
PN WO200281670-A1.
PD 17-OCT-2002.
PF 04-APR-2002; 2002WO-US10786.
PR 06-APR-2001; 2001US-282036P.
PA (LEXI-) LEXICON GENETICS INC.
XX
PI Turner CA, Mathur B, Friddle CJ;
XX
DR WPI; 2003-058538/05.
DR N-PSDB; ABV74557.
XX
New human kinase proteins useful for diagnosis, drug screening, and
PT clinical trial monitoring, treatment of disorders and diseases, and
PT cosmetic and nutritional applications -
XX
PS Claim 5; Page 40-41; 47pp; English.
XX
CC The present sequence is a novel human kinase. The genomic locus encoding
CC the kinase is thought to be on human chromosome 19. The kinase and its
CC coding sequence are useful for diagnosis, drug screening, clinical trial
CC monitoring, treatment of disorders and diseases, and cosmetic and
CC nutritional applications.
XX
SQ Sequence 778 AA;

Query Match 79.8%; Score 2545.5; DB 24; Length 778;
Best Local Similarity 75.6%; Pred. No. 1.7e-180;
Matches 495; Conservative 51; Mismatches 44; Indels 65; Gaps 7;

Qy 1 MKVEREIALKLTIEHPHVLKLDHVDYENKKYLVLEHVS GGELFDYLVKKGRLLTPKEARK 60
Db 76 MKVEREIALKLTIEHPHVLKLDHVDYENKKYLVLEHVS GGELFDYLVKKGRLLTPKEARK 135
Qy 61 FFRQISALDFCHSHSICHRLDKPENLLDEKNRIADFGMASLQVGDLSLETSCGSPH 120
Db 136 FFRQIVSALDFCHSHSICHRLDKPENLLDEKNRIADFGMASLQVGDLSLETSCGSPH 195
Qy 121 YACPEVIREKYGDKADYVWSCGVILFALLVGLPDDNNLRQLLEKVKRGVFNHPHIP 180
Db 196 YACPEVIRKEKYGDKADYVWSCGVILFALLVGLPDDNNLRQLLEKVKRGVFNHPHIP 255
Qy 181 PQOSLLRGMIEVDAAARLLTLEHIOKHIWYIGGKNEPEP-EQIP-RKQVIRSLPSLEDI 238
XX
```

```
Db 256 PDCOSLLRGMIEVEPEKRLSLEQIKHPWYLGKHEDPCLEPAPGRVAMRSLPSNGEL 315
Qy 239 DPVDLSMHSILGCFRDRNKLQDLLSEENQEKMIYFLLLDKERYPSQEDDLPFRNEI 298
Db 316 DPVDLESMAISLGCFRDRERLHRELRSSEENQEKMIYFLLLDKERYPSCEDQLPRNDV 375
Qy 299 DPPKRVDS PMLNRHGRKRRPERKSMELSVTD----GGSPVAPRAIEMAHQGQRSISG 355
Db 376 DPPKRVDS PMLSRHGRKRRPERKSMELSVTDAGGGSPVPTRRALEMAHQSORSRVSG 435
Qy 356 ASSGLSTPLSSPR-----VTPHP 374
Db 436 ASTGLSSPLSSPRSPVPSFSPSPGAGDEARGGSPSTQTLPSPRPGGGAGEQPPPP 495
Qy 375 SPRGSPPLTPKG-----TPVHTPKESPAGTNPPTPPSSP--SVGVPMWRRLNS 421
Db 496 SARSTPLPGPPGSPRSSGGTPLHSLPLTPRASPTGTPTTPPPSPGGGSGVGAARSLNS 555
Qy 422 IKNSFLGSPRFRKLOVPTPEMSNLTPESSPELAKKSMFGNEISLEKEQIFVVIKDK 481
Db 556 IRNSFLGSPRFRKMQVPTAEEMSLTPESPSELAKRSWFGNFIISLDKEQIFLVLDK 615
Qy 482 PLSSIKADIVHAFSLIPSLSHSVISQTSFRAEYKATGPAVFQKPVKQVDITYEGGEA 541
Db 616 PLSSIKADIVHAFSLIPSLSHSVISQTSFRAEYKASGGPSVFQKPVQVDISSSEGEPEP 675
Qy 542 QKE-----NGIYSVTFTLLSGPSRRFRKRVVETIQALLSTHDPAAQHLSDTTN 590
Db 676 SPRDGGGGGIYSVTFTLLSGPSRRFRKRVVETIQALLSTHDPQSVQALADEKN 730

RESULT 9
AAU03517
ID AAU03517 standard; Protein; 794 AA.
XX
AC AAU03517;
XX
DT 12-SEP-2001 (first entry)
DE Human protein kinase #17.
XX
KW Human; protein kinase; PTK; STK; cancer; cardiovascular disease;
KW metabolic disorder; immune related disease; neurological disorder;
KW neurodegenerative disorder; inflammatory disorder; infectious disease;
KW reproductive disorder.
XX
OS Homo sapiens.
XX
PN WO200138503-A2.
XX
PD 31-MAY-2001.
XX
PF 22-NOV-2000; 2000WO-US32085.
XX
PR 24-NOV-1999; 99US-0167482.
XX
PA (SUGE-) SUGEN INC.
XX
PI Plowman GD, Whyte D, Manning G, Sudarsanam S, Martinez R;
XX
XX Flanagan P, Clary D;
XX
DR WPI; 2001-343950/36.
XX
DR N-PSDB; AAS06717.
XX
Nucleic acids encoding human kinase polypeptides, useful for preventing
PT diagnosing and/or treating e.g. cancer, immune, cardiovascular and
PT neuronal-associated diseases, and microbial infections -
XX
PS Claim 7; Figure 2; 433pp; English.
XX
CC AAU03501-AAU03557 represent novel human protein kinases #1-57. The
CC novel protein kinases have been identified as members of the tyrosine
CC or serine/threonine kinase (PTK and STK) families. The polynucleotides
```

encoding protein kinases and the polypeptides may be used in the prevention, diagnosis and treatment of diseases associated with inappropriate kinase expression. For example, they may be used to treat cancers (especially cancers of haematopoietic origin), cardiovascular disease (e.g. atherosclerosis), metabolic disorders (e.g. diabetes), immune related diseases (e.g. rheumatoid arthritis), neurological disorders (e.g. schizophrenia), neurodegenerative disorders (e.g. Parkinson's disease), inflammatory disorders (e.g. asthma), infectious disease (e.g. HIV) and reproductive disorders (e.g. infertility). Additionally, polynucleotides encoding protein kinases may be used for gene therapy and as DNA probes in diagnostic assays. The protein kinase polypeptides may be used as antigens in the production of antibodies against the protein kinases and in assays to identify modulators of protein kinase expression and activity.

SQ	Sequence	794 AA;
	Query Match	79.8%; Score 2545.5; DB 22; Length 794;
	Best Local Similarity	75.6%; Pred. No. 1.8e-180;
	Matches 495:	Conservative 51; Mismatches 44; Indels 65; Gaps 7

Qy	1	MKVREIEAILKLIIEHPHVKLHDVYENKCYLYLVLEHVSGGELFDYLVKKGRLLTPKEARK	60
Db	92	MKVREIEAILKLIIEHPHVKLHDVYENKCYLYLVLEHVSGGELFDYLVKKGRLLTPKEARK	151
Qy	61	FFRQIISALDFCHSHSICHRDLKPENLILDEKNNIRIADFQMASLOVGDLSLETSCGSPH	120
Db	152	FFRQIVSALDFCHSYVICHRDLKPENLILDEKNNIRIADFQMASLOVGDLSLETSCGSPH	211
Qy	121	YACPEVIRGEKYDGRKADVWSGGIVLFPALLVGAALPFDDDLNLRQLLEKVKRGVFMHPHFTP	180
Db	212	YACPEVIRGEKYDGRADVWSGGIVLFPALLVGAALPFDDDLNLRQLLEKVKRGVFMHPHFTP	271
Qy	181	PDQSQLLRMGIEVDAARLLTLEHQIKHWYIGGKNEPEP-EQIIP-RKVOIRSLPSLEDI	238
Db	272	PDQSQLLRMGIEVEPEKLSLEEQIKHPWYLGKGHEFDDPCLEFAPGRVAMRSLPENGEL	331
Qy	239	DPDVLDSMHSLGCFDRNKLQDLLSEBENQEKMIYFLLLDKREKYPDSQEDDLPPRNEI	298
Db	332	DPDVLDSMASLGCPRDRERLHRELSEBENQEKMIYLLLDKREKYPSCDQDLPPRNDV	391
Qy	299	DPERKRVDSPLNLRHCKRPERKSNVLSVTD--GGSPVAPRAIEMAHGORSRSISG	355
Db	392	DPERKRVDSPLNLSRHGKRPKRSMEVLSITDAGGGGSPVPTRRALAEAHGORSRSVSG	451
Qy	356	ASSGLSTSPLSLSPR-----VTTPH	374
Db	452	ASTGLSSPLSSPRSPVFSFPEPCAGDEARGGGSPSTKQTLPSPRGCGAGGEQPPPP	511
Qy	375	SPFGSPLPATPKG-----TPVHTPKSPAGTNPPTPPSP--SVGGVFWRAALNS	421
Db	512	SARSTPLPGSPGSPRSSGGTPLHSPSLHTPRASPTGTPTGTPPPSPGGVGGAARSLNS	571
Qy	422	IKNSFLGSPFRHRRKLOVPTPEMSNLTPESPSPELAKSNWFGNFIISLEKEEQIFVWIKDK	481
Db	572	IRNSFLGSPFRHRRKQVPTAEKMSLTPESPSPELAKSNWFGNFIISLDKEEQIFLVLIKDK	631
Qy	482	PLSSIKADIVHAPLSIPSLSHSVISQTSFRAEYKATGGPAVQKPKVQVDITVYTGGEA	541
Db	632	PLSSIKADIVHAFSLPSLSHSLVLSQTSFRAEYKASGGPSVQKPVAFQVDISSESQEP	691
Qy	542	QKE-----NGIYSVTFTLLSGPSRRFRKRVVETIQALLSTHDPDPAQHLSDDITN	590
Db	692	SPRDRGGGGGIYSVTFTLLISGPSRRFRKRVVETIQALLSTHDPQSVQALADEKN	746
RESULT 10			
AAE16271			
ID	AAE16271 standard; Protein; 794 AA.		
XC			
AC	AAE16271;		
XX			
DT	26-MAR-2002 (first entry)		

RESULT 10
AAE16271
ID AAE1
XX
AC AAE1
XX
DT 26-M

XX	Human kinase PKIN-17 protein.
DE	Human; kinase; PKIN-17; cancer; leukaemia; adenocarcinoma; osteoporosis;
XX	immune disorder; atherosclerosis; Crohn's disease; Hodgkin's disease;
XX	Acquired Immune Deficiency Syndrome; AIDS; Addison's disease; anaemia;
XX	allergy; asthma; adult respiratory distress syndrome; multiple sclerosis;
XX	autoimmune thyroiditis; bronchitis; diabetes mellitus; osteoarthritis;
XX	Good pasture's syndrome; Graves' disease; pancreatitis; psoriasis;
XX	rheumatoid arthritis; ulcerative colitis; cirrhosis; Cushing's syndrome;
XX	hepatitis; hypothyroidism; cerebral palsy; cataract; angina pectoris;
XX	cardiovascular disease; hypertension; vasculitis; myocarditis; obesity;
XX	congestive heart failure; ischaemic heart disease; lung tumour; gout;
XX	fatty liver; Niemann-Pick's disease; gene therapy.
XX	
OS	Homo sapiens.
XX	
FF	Key
FF	Location/Qualifiers
FF	Domain
FF	50..301
FF	/note= "Eukaryotic protein kinase domain"
FF	Domain
FF	51..292
FF	/label= Protein_kinase_domain
FF	Domain
FF	52..292
FF	/label= Protein_kinase_domain
FF	Domain
FF	71..292
FF	/label= Protein_kinase_domain
XX	
PN	WO200196547-A2.
XX	
XX	
PD	20-DEC-2001.
XX	
FF	14-JUN-2001; 2001WO-US19444.
XX	
PR	15-JUN-2000; 2000US-212073P.
PR	23-JUN-2000; 2000US-213467P.
PR	30-JUN-2000; 2000US-215651P.
PR	07-JUL-2000; 2000US-216603P.
PR	13-JUL-2000; 2000US-218372P.
PR	25-AUG-2000; 2000US-228056P.
XX	
PA	(INCYTE) INCYTE GENOMICS INC.
XX	
PI	Yue H, Ial P, Bandman O, Borowsky ML, Au-Young J, Lu Y;
PI	Gandhi AR, Tribouley CM, Wallia NK, Yao MG, Lu DAM, Greenwald SR;
PI	Ramkumar J, Griffin JA, Kearney L, Burford N, Nguyen DB, Tang YT;
PI	Baughn MR, He A, Thornton M, Hafalia A, Patterson C, Gururajan R;
PI	Lo TP, Khan F, Recipon SA, Azimzai Y, Policky JL, Ding L;
PI	Grether M, Elliott VS, Thangavelu K, Batra S, Ison CH;
XX	
DR	WPI; 2002-090207/12.
DR	N-PSDB; AAD26464.
XX	
PT	New polypeptides, useful for diagnosing, treating or preventing
PT	disorders of growth and development, cardiovascular and lipid, and
PT	diseases such as cancer, comprise human kinase polypeptides -
XX	
XX	Claim 1; Page 159-161; 197pp; English.
XX	
CC	The invention relates to human kinase PKIN proteins and their
CC	corresponding cDNAs. A composition containing PKIN agonist is useful for
CC	treating a disease or condition associated with decreased expression of
CC	PKIN and a composition comprising PKIN antagonist is useful for treating
CC	a disease or condition associated with overexpression of PKIN. The
CC	disorders include cancer (leukaemia, adenocarcinoma, lymphoma, melanoma,
CC	myeloma, sarcoma, teratocarcinoma, Hodgkin's disease); immune disorder
CC	(Acquired Immune Deficiency Syndrome (AIDS), asthma, Addison's disease,
CC	atherosclerosis, anaemia, allergies, adult respiratory distress syndrome,
CC	autoimmune thyroiditis, gout, bronchitis, Crohn's disease, diabetes
CC	mellitus, multiple sclerosis, Good pasture's syndrome, Graves' disease,
CC	osteoarthritis, osteoporosis, pancreatitis, psoriasis, Reiter's syndrome,
CC	rheumatoid arthritis, Sjogren's syndrome, uveitis, ulcerative colitis,
CC	bacterial, parasitic, fungal, viral, protozoal and helminthic infections,
CC	growth and development disorders (arteriosclerosis, cirrhosis, hepatitis,

CC Cushing's syndrome, hypothyroidism, cerebral palsy, cataracts); cardio
CC vascular disease (arteriovenous fistula, hypertension, vasculitis,
CC aneurysms, congestive heart failure, angina pectoris, myocarditis,
CC ischaemic heart disease, chronic bronchitis, lung tumours); lipid
CC disorder (fatty liver, Fabry's disease, Niemann-Pick's disease,
CC hypcholesterolaemia, obesity). PKIN DNA is useful for assessing
CC toxicity of a test compound and in gene therapy. The present sequence
CC is human PKIN-17 protein.
XX
SQ Sequence 794 AA;

Query Match 79.8%; Score 2545.5; DB 23; Length 794;
Best Local Similarity 75.6%; Pred. No. 1.8e-180;
Matches 495; Conservative 51; Mismatches 44; Indels 65; Gaps 7;

QY 1 MKVEREIAILKLEHSHVILKLDVYENKYLVLVLEHVSGLGELFDYLVKGRGLTPKEARK 60
DB 92 MKVEREIAILKLEHSHVILKLDVYENKYLVLVLEHVSGLGELFDYLVKGRGLTPKEARK 151

QY 61 FFRQITSLDCHSHSICHRLDKPENLLDEKNIRIADFGMASLQVDSLETSCGSPH 120
DB 152 FFRQIVSALDFCHSYISICHRLDKPENLLDEKNIRIADFGMASLQVDSLETSCGSPH 211

QY 121 YACPEVIRGKYGKADVMSGCVILFALLVGLPFDNRLQLLEKVRGVFHPHFIP 180
DB 212 YACPEVIRGKYGKADVMSGCVILFALLVGLPFDNRLQLLEKVRGVFHPHFIP 271

QY 181 PDCQSLRGMIIEVDAAARLTLEHIOKHIVYIGGKNEPEP-EQIP-RKQVIRSLPSLEDI 238
DB 272 PDCQSLRGMIIEVEPEKLSLEIQKHVPYLGKHEPDCLEPAPGRVAVRSLPSNGEL 331

QY 239 DPVLDLSMHSGCFDRNKLQDLSEENQEKMIYFLLDKRYPSQDEDELPNNEI 298
DB 332 DPVLESMSLGCPRDRRLHRELSEENQEKMIYFLLDKRYPSQDEDELPNNDV 391

QY 299 DPPRKVDSPMLNRHGRKRRPERKSNVLSVTD---GGSPVPARRAIEAQAQGRSRSISG 355
DB 392 DPPRKVDSPMLNRHGRKRRPERKSNVLSVTD---GGSPVPARRAIEAQAQGRSRSISG 451

QY 356 ASSGLSTSPSSPR-----VTPH 374
DB 452 ASTGLSSPLSSPRSPVFSFPEPGAGDEARGGSPSTKTQTLPSRPGGAGEQPPPP 511

QY 375 SPRGSPLTPKG-----TPVHTPKSPAGTNPPTPPSP--SVGGVPWRRLNS 421
DB 512 SARSTPLPGPGSPRSSGCTPLHSPHTPRASPTGPTGPTTPPSPGGVGGAAWRSLNS 571

QY 422 IKNSPLGSPRHRRLKQVPTPEMSNLTPESSPELAKKSWFGNISLEKEROIFVVKDK 481
DB 572 IRNSPLGSPRHRRLKQVPTPEMSNLTPESSPELAKKSWFGNISLEKEROIFVVKDK 631

QY 482 PLSSIKADIVHAFSLIPSLHSVISOTSFRABYKATGGPAVQKPVQVDITVTEGGEA 541
DB 632 PLSSIKADIVHAFSLIPSLHSVISOTSFRABYKASGGSPVQKPVQVDISSEGEPEP 691

QY 542 QKE-----NGIYSVTFTLLSGPSRRFRVVTETIQALLSTHDPAAQHLSDTTN 590
DB 692 SPRDGGSGGGIYSVTFTLLSGPSRRFRVVTETIQALLSTHDPQSVQALADEKN 746

RESULT 11
AAO16604
ID AAO16604 standard; protein; 754 AA.
XX
AC AAO16604;
XX
DT 08-MAY-2003 (first entry)
XX
DE Human cell cycle-regulatory factor Cdr2.
XX
KW Human; cell cycle-regulatory factor; Cdr2; kinase; proliferative disease;
KW anticancer agent; wound-healing drug.
XX

OS Homo sapiens.
FN WO200299110-A1.
XX
PD 12-DEC-2002.
XX
PF 03-JUN-2002; 2002WO-JP05411.
XX
PR 04-JUN-2001; 2001JP-0168792.
XX
PA (TAIH) TAIHO PHARM CO LTD.
PA (NAXA/) NAKANISHI M.
XX
PI Nakanishi M;
XX
DR WPI; 2003-156857/15.
DR N-PSDB; AAL51889.
XX
PT Cell cycle-regulatory factor Cdr2 with kinase activity and encoded
PT gene, applicable in diagnosis of and screening drugs for proliferative
PT diseases e.g. anticancer agents and wound-healing drugs -
XX
PS Claim 1; Fig 2 A-C; 63pp; Japanese.
XX
CC The invention comprises the amino acid and coding sequence of the human
CC cell cycle-regulatory factor Cdr2 (with kinase activity). The DNA and
CC protein sequences of the invention are useful in diagnosing and screening
CC drugs for proliferative diseases (e.g. anticancer agents and wound-
CC healing drugs). The present amino acid sequence represents the human cell
CC cycle-regulatory factor Cdr2 protein.
XX
SQ Sequence 754 AA;

Query Match 78.9%; Score 2514.5; DB 24; Length 754;
Best Local Similarity 77.5%; Pred. No. 3.3e-178;
Matches 489; Conservative 51; Mismatches 50; Indels 41; Gaps 8;

QY 1 MKVEREIAILKLEHSHVILKLDVYENKYLVLVLEHVSGLGELFDYLVKGRGLTPKEARK 60
DB 76 MKVEREIAILKLEHSHVILKLDVYENKYLVLVLEHVSGLGELFDYLVKGRGLTPKEARK 135

QY 61 FFRQITSLDCHSHSICHRLDKPENLLDEKNIRIADFGMASLQVDSLETSCGSPH 120
DB 136 FFRQIVVACGLCHSYISICHRLDKPENLLDEKNIRIADFGMASLQVDSLETSCGSPH 195

QY 121 YACPEVIRGKYGKADVMSGCVILFALLVGLPFDNRLQLLEKVRGVFHPHFIP 180
DB 196 YACPEVIRGKYGKADVMSGCVILFALLVGLPFDNRLQLLEKVRGVFHPHFIP 255

QY 181 PDCQSLRGMIIEVDAAARLTLEHIOKHIVYIGGKNEPEP-EQIP-RKQVIRSLPSLEDI 238
DB 256 PDCQSLRGMIIEVEPEKLSLEIQKHVPYLGKHEPDCLEPAPGRVAVRSLPSNGEL 315

QY 239 DPVLDLSMHSGCFDRNKLQDLSEENQEKMIYFLLDKRYPSQDEDELPNNEI 298
DB 316 DPVLESMSLGCPRDRRLHRELSEENQEKMIYFLLDKRYPSQDEDELPNNDV 375

QY 299 DPPRKVDSPMLNRHGRKRRPERKSNVLSVTD---GGSPVPARRAIEAQAQGRSRSISG 355
DB 376 DPPRKVDSPMLNRHGRKRRPERKSNVLSVTD---GGSPVPARRAIEAQAQGRSRSISG 435

QY 356 ASSGLSTSPSSPR-----VTPH 387
DB 436 ASTGLSSPLSSPRSPVFSFPEPGAGDEARGGSPSTKTQTLPGPGSPRSSGGTPLHS 495

QY 388 PVHTPKSPAGTNPPTPPSP--SVGGVPWRRLNSIKNSFLSGSPRHRRLKQVPTPEEM 445
DB 496 PLHTPRASPTGPTGPTTPPSPGGVGGAAWRSLNSIRNSFLSGSPRHRRLKQVPTPEEM 555

QY 446 SNLTPESSPELAKKSWFGNISLEKEROIFVVKDKPLSSIKADIVHAFSLPSLSHVI 505
DB 556 SSLETPESSPELAKKSWFGNISLEKEROIFVVKDKPLSSIKADIVHAFSLPSLSHVI 615

Qy 506 SOTSFRAYKATGPAVQKPVKFOVDITYTEGGAQKE-----NGIYSVTFTLLSGPS 559
Db 616 SOTSFRAYKASGGSPVFQKPVRFQVDISSSEGEPSPRDRGSGGGGIYSVTFTLLSGPS 675
Qy 560 RRFKRVVETIQALLSTHDPPAAQHLSDTTN 590
Db 676 RRFKRVVETIQALLSTHDQPSVQALADEKN 706
RESULT 12
ABB98744
ID ABB98744 standard; Protein; 762 AA.
XX AC ABB98744;
XX DT 20-JAN-2003 (first entry)
XX DE Human kinase #2.
XX KW Human; kinase; chromosome 19.
XX OS Homo sapiens.
XX WO20020281670-A1.
XX PD 17-OCT-2002.
XX PF 04-APR-2002; 2002WO-US10786.
XX PR 06-APR-2001; 2001US-282036P.
XX PA (LEXI-) LEXICON GENETICS INC.
XX PI Turner CA, Mathur B, Friddle CJ;
XX WPI; 2003-058538/05.
XX DR N-PSDB; ABV74558.
XX PT New human kinase proteins useful for diagnosis, drug screening, and
XX PT clinical trial monitoring, treatment of disorders and diseases, and
XX PT cosmetic and nutritional applications -
XX PS Claim 5; Page 42-43; 47pp; English.
XX CC The present sequence is a novel human kinase. The genomic locus encoding
CC the kinase is thought to be on human chromosome 19. The kinase and its
CC coding sequence are useful for diagnosis, drug screening, clinical trial
CC monitoring, treatment of disorders and diseases, and cosmetic and
CC nutritional applications.
SQ Sequence 762 AA;
Query Match 74.9%; Score 2386.5; DB 24; Length 762;
Best Local Similarity 74.1%; Pred. No. 1.1e-168;
Matches 464; Conservative 52; Mismatches 45; Indels 65; Gaps 7;
Qy 30 YLYLVLEHVSQGELEFDLVKXGRLTPKEARFFQIISALDFCHSHSICHRLDKPENLL 89
Db 89 FRYLVLEHVSQGELEFDLVKXGRLTPKEARFFQIISALDFCHSHSICHRLDKPENLL 148
Qy 90 DEKNIRIADFGMASLQVDSLLTSCGSPHYACPEVIRGEKYDOKKADVWSCGVILPAL 149
Db 149 DEKNIRIADFGMASLQVDSLLTSCGSPHYACPEVIRGEKYDOKKADVWSCGVILPAL 208
Qy 150 LVGALPFDDNLRQLLEKVRGVFHMHPFIPDCOSLLRGMTEVDAAARLTLEHIQHIW 209
Db 209 LVGALPFDDNLRQLLEKVRGVFHMHPFIPDCOSLLRGMTEVDAAARLTLEHIQHIW 268
Qy 210 YTGKNEPEP-EQPIP-RKVQIRSLPSLEIDIPDVLDSNHSISGCFRDRNKLQDLSSEE 267
Db 269 YLGKHEPDPCLEPAPGRVAMRSLPSNGELDPDVLDSNHSISGCFRDRNKLQDLSSEE 328
Qy 268 NOEKMIYFLLDRKERYPSQEDLPPRNEIDPPKRVDSPLNRHGKRRPERKSMVLS 327

Db 329 NQEKMIYFLLDRKERYPSQEDLPPRNDVDPKRVDSPLNRHGKRRPERKSMVLS 388
Qy 328 VTD---GGSPVAPRAIEMAHQGRSRTSGASSGLSTPLSSPR----- 369
Db 389 ITDAGGGSPVPTRRALEMAQHSQSRSVSGASTGLSSPLSSPRSPVFSPEPCAGDE 448
Qy 370 -----VTPHSPRGSPLFTPKG-----TPVHTP 392
Db 449 ARGGSPTSKTQLPSRGRPGGAGCQPPPPSARSTPLFGPPGSPRSSGGTPLHSLHTP 508
Qy 393 KESPACTNPPTPPSP--SVGVFWRARLNSKNSFLGSPRFRHRKQLQVPTPEMSNLTP 450
Db 509 RASPTGCTGTTTPPSPGGVGGAWRSLNSIRNSFLGSPRFRHRKQVPTAEWMSLTP 568
Qy 451 ESSPELAKKSWFGNFTSLERQIFVWIKDKPLSSIKADIVHAFSLPSLSHVSISQTSF 510
Db 569 ESSPELAKKSWFGNFTSLERQIFVWIKDKPLSSIKADIVHAFSLPSLSHVSISQTSF 628
Qy 511 RAEYKATGPAVQKPVKFOVDITYTEGGAQKE-----NGIYSVTFTLLSGPSRRFKR 564
Db 629 RAEYKASGGSPVFQKPVRFQVDISSSEGEPSPRDRGSGGGGIYSVTFTLLSGPSRRFKR 688
Qy 565 VVETIQALLSTHDPPAAQHLSDTTN 590
Db 689 VVETIQALLSTHDQPSVQALADEKN 714
RESULT 13
ABB62061
ID ABB62061 standard; Protein; 851 AA.
XX AC ABB62061;
XX DT 26-MAR-2002 (first entry)
XX DE Drosophila melanogaster polypeptide SEQ ID NO 12975.
XX KW Drosophila; developmental biology; cell signalling; insecticide;
XX KW pharmaceutical.
XX OS Drosophila melanogaster.
XX WO200171042-A2.
XX PD 27-SEP-2001.
XX PF 23-MAR-2001; 2001WO-US09231.
XX PR 23-MAR-2000; 2000US-191637P.
XX PR 11-JUL-2000; 2000US-0614150.
XX PA (PEKE) PE CORP NY.
XX PI Venter JC, Adams M, Li PWD, Myers EW;
XX WPI; 2001-656860/75.
XX DR N-PSDB; ABL06164.
XX PT New isolated nucleic acid detection reagent for detecting 1000 or more
XX PT genes from Drosophila and for elucidating cell signalling and cell-cell
XX PT interactions -
XX PS Disclosure; SEQ ID NO 12975; 21pp + Sequence Listing; English.
XX CC The invention relates to an isolated nucleic acid detection reagent
CC capable of detecting 1000 or more genes from Drosophila. The invention is
CC useful in developmental biology and in elucidating cell signalling and
CC cell-cell interactions in higher eukaryotes for the development of
CC insecticides, therapeutics and pharmaceutical drugs. The invention
CC discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA
CC sequences (ABL01840-ABL16175) and the encoded proteins
CC (ABB57737-ABB72072).

CC screening techniques. The present sequence represents a novel human
CC polypeptide of the invention.

XX
SQ Sequence 301 AA;
Query Match 32.5%; Score 1036.5; DB 22; Length 301;
Best Local Similarity 79.4%; Pred. No. 7.2e-69;
Matches 197; Conservative 10; Mismatches 6; Indels 35; Gaps 1;
Qy 1 MKVERETAILKLIHPHVLKLDVYENKKYL----- 31
Db 50 MKVERETAILKLIHPHVLKLDVYENKKYPPDELTSGPSMLAQVSPHGKLSARRSWDL 109
Qy 32 -----YLVLEHVSGLGELFDYLVKKGRLTPKARKFFRQIISALDFCHSHSICHRLDKPE 85
Db 110 LSGFPRIYLVLEHVSGLGELFDYLVKKGRLTPKARKFFRQIISALDFCHSYICHRLDKPE 169
Qy 86 NLLDEKNNIRIADFGMASLQVDSLSLETSCGSPHYACPEVIRGEKYDGRKADVMSCGVI 145
Db 170 NLLDEKNNIRIADFGMASLQVDSLSLETSCGSPHYACPEVIRGEKYDGRADVMSCGVI 229
Qy 146 LFALLVGLPDDNNLRQLLEKVKRGVFMHPHFIPDDCQSLRGMIWDAAARLTLEHIQ 205
Db 230 LFALLVGLPDDNNLRQLLEKVKRGVFMHPHFIPDDCQSLRGMIWDAAARLTLEHIQ 289
Qy 206 KHIWYIGG 213
Db 290 KHPWYLG 297

RESULT 15
AAG70854
ID AAG70854 standard; Protein; 1349 AA.

AC AAG70854;

DT 27-JUL-2001 (first entry)

DE C albicans apoptosis associated protein #34.

XX Yeast; fungus; apoptosis; infection; proliferative disease;
KW vaccine; autoimmune disease; ischaemia; neurodegeneration.

XX Candida albicans.

XX WO200102550-A2.

XX 11-JAN-2001.

XX 03-JUL-2000; 2000WO-BE00077.

XX 01-JUL-1999; 99EP-0870141.

XX (JANC) JANSSEN PHARM NV.

XX Contreras RH, De Backer MD, Luyten WHML, Malcorps IKL;
PI Nelissen BJM, Reekmans RJ;

XX WPI: 2001-367042/38.

XX N-PSDB; AAH29890.

XX Yeast and fungal nucleic acids encoding proteins involved in a pathway
PT leading to programmed cell death, useful for treating proliferative
PT disorders, yeast and fungal infections, or for preventing apoptosis in
PT certain diseases -

PS Claim 24; Fig 2; 218pp; English.

XX The present invention provides the protein and coding sequences of a
CC number of apoptosis associated proteins from the yeast *Saccharomyces*
CC *cerevisiae* and the fungus *Candida albicans*. These can be used to identify
CC treatments for fungal and yeast infections, for proliferative diseases
CC and for apoptosis related diseases such as autoimmune diseases, ischaemia

CC and neurodegeneration. The present sequence is one of the C. albicans
CC proteins of the invention.

XX
SQ Sequence 1349 AA;

Query Match 24.0%; Score 766.5; DB 22; Length 1349;
Best Local Similarity 33.5%; Pred. No. 6e-48;
Matches 211; Conservative 97; Mismatches 216; Indels 105; Gaps 20;
Qy 3 VERETAILKLIHPHVLKLDVYENKKYLVLVLEHVSGLGELFDYLVKKGRLTPKARKFF 62
Db 80 IEREITIMKLTNPVRLYDWMETSKALYLVLEVVEGELFDLLVERGLPDEVAIKFF 139
Qy 63 RQIISALDFCHSHSICHRLDKPENLILDEKNNIRIADFGMASLQVDSLSLETSCGSPHYA 122
Db 140 RQIILGTAYCHALGICHRLDKPENLILDSQLNVKLADFGMAALESNGKLETSCGSPHYA 199
Qy 123 CEPVIRGEKYDGRKADVMSCGVILFALLVGLPDDNNLRQLLEKVKRGVFMHPHFIP 181
Db 200 APEIVSGLKYHGAASDVMSCGVILFALLVGLPDDNNLRQLLEKVKRGVFMHPFVDEVR 259
Qy 182 DCQSLRGMIWDAAARLTLEHIQKH---IWIIGKNEPEPEQPIPR-KVQIRSLPSLED 237
Db 260 EARDLIRLMLEVDPMRRISTEKILRHPLTKYPMNSNEDLISEKSLPHPTGYKSLGSVRN 319
Qy 238 IDPDVLDMSHSLGCFRDRNKLQDLLESEENOEKMIYFLLLDKERYPSQEDDLPPRNE 297
Db 320 IDKQILSNLTILWDRPEBEIVDCLLKDGSPKFTYALLM---RYKHNQDDN---TNN 372
Qy 298 IDPPRKRVDSFMLNRHGKR-----RPERKSMVLSVTDGSPVPARRAIEMAHQGORSRS 352
Db 373 NSPKKSTSFNNKVVRSKYSKSLNGTTPRRKRASHISVS-----RPTSFQYKSNP 420
Qy 353 ISGASSGLSTSPSSPRVTPHPSRGSPLTPKGTVPVHTPKESPAG--TNPPTPPSSPSV 410
Db 421 GAGATANRNSVARHSVASSANNSPRKSPYKSPYRSPYKSPYKSPYKSPYKSPYKSPYKSP 477
Qy 411 GGVPWRARLNSIKNSFLGSP-RFHRRLKLVPTPEMSNLT-PESSPELAKKSWFGNFISL 468
Db 478 ---PYGRSNS-QRFENEPLKAKPRNIYNEIVDAQSNFSLPPLPP-----SL 522
Qy 469 EKEEQIFVVKDKPLSSIKADIVHAFSLPSLSHVSISQTSFRAEYKATGCPAVFQKPVK 528
Db 523 PSKDSRYMI--DEP-----NOPLOQPALSQ-----VPENPVDESP-- 557
Qy 529 FQVDITVTGGEAKENGIV-----SVTFLLSGPSRRFRKRVETIOA 571
Db 558 ---DLMQSAKISSGKNSIIGKNNNSNKNRMSKRSIRASMTTGLKR-----NSITM 608
Qy 572 QLLSTHDPAAQHLSDTTNCMEMTGLRLS 600
Db 609 KLLSTY---AKLSGDDDDWEYMDKQTKRTS 634

Search completed: November 26, 2003, 11:59:56
Job time : 38.6897 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 26, 2003, 11:58:31 ; Search time 14.2947 Seconds
(without alignments)
1799.622 Million cell updates/sec

Title: US-10-054-579-4
Perfect score: 3188
Sequence: 1 MKVERETAILKLEHPPHVLK.....TNCMEMTGRLSKCGIIPKS 608

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 328717 seqs, 42310858 residues

Total number of hits satisfying chosen parameters: 328717

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents AA:*
1: /cgn2_6/ptodata/1/1aa/5A_COMB.pep:*
2: /cgn2_6/ptodata/1/1aa/5B_COMB.pep:*
3: /cgn2_6/ptodata/1/1aa/6A_COMB.pep:*
4: /cgn2_6/ptodata/1/1aa/6B_COMB.pep:*
5: /cgn2_6/ptodata/1/1aa/PTCUS_COMB.pep:*
6: /cgn2_6/ptodata/1/1aa/backfiles.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	3188	100.0	668	US-09-930-181-2	Sequence 2, Appli
2	3028	95.0	585	US-09-930-181-4	Sequence 4, Appli
3	3015	94.6	603	US-09-930-181-17	Sequence 17, Appl
4	688.5	21.6	745	US-09-523-849-36	Sequence 36, Appl
5	685.5	21.5	724	US-09-984-890-2	Sequence 2, Appli
6	684.5	21.5	722	US-09-984-890-4	Sequence 4, Appli
7	682.5	21.4	722	US-08-817-832B-32	Sequence 32, Appl
8	680	21.3	149	US-09-930-181-18	Sequence 18, Appl
9	663.5	20.8	729	US-08-677-298-2	Sequence 2, Appli
10	663.5	20.8	729	US-09-523-849-33	Sequence 33, Appl
11	657.5	20.6	793	US-09-523-849-32	Sequence 32, Appl
12	642	20.1	779	US-08-817-832B-31	Sequence 31, Appl
13	638	20.0	776	US-09-523-849-34	Sequence 34, Appl
14	636	19.9	552	US-08-557-006C-40	Sequence 40, Appl
15	621.5	19.5	345	US-09-101-146-1	Sequence 1, Appli
16	616	19.3	633	US-08-557-006C-43	Sequence 43, Appl
17	606	19.0	257	US-09-101-146-6	Sequence 6, Appli
18	581.5	18.2	257	US-07-857-224B-25	Sequence 25, Appl
19	579	18.2	604	US-09-523-849-35	Sequence 35, Appl
20	553.5	17.4	631	US-09-579-664B-11	Sequence 11, Appl
21	461.5	14.5	252	US-07-857-224B-26	Sequence 26, Appl
22	458.5	14.4	353	US-08-688-988-31	Sequence 31, Appl
23	458.5	14.4	363	US-08-688-988-30	Sequence 30, Appl
24	451	14.1	260	US-07-857-224B-27	Sequence 27, Appl
25	449.5	14.1	354	US-08-688-988-29	Sequence 29, Appl
26	445.5	14.0	351	US-08-688-988-28	Sequence 28, Appl
27	442.5	13.9	339	US-08-688-988-33	Sequence 33, Appl

28	440	13.8	359	3	US-08-688-988-32	Sequence 32, Appl
29	439	13.6	260	2	US-07-857-224B-28	Sequence 28, Appl
30	434	13.6	355	4	US-09-579-664B-10	Sequence 10, Appl
31	433	13.6	370	2	US-08-878-989-19	Sequence 19, Appl
32	433	13.6	370	3	US-09-272-796-19	Sequence 19, Appl
33	433	13.6	370	4	US-09-457-040B-31	Sequence 31, Appl
34	432.5	13.6	1037	4	US-09-428-711A-21	Sequence 21, Appl
35	429.5	13.5	556	4	US-09-800-960-4	Sequence 4, Appli
36	427	13.4	290	4	US-09-734-673-4	Sequence 4, Appli
37	427	13.4	334	4	US-09-523-849-31	Sequence 31, Appl
38	425	13.3	339	3	US-08-688-988-2	Sequence 2, Appli
39	423	13.3	565	4	US-09-800-960-2	Sequence 2, Appli
40	420.5	13.2	436	4	US-09-734-673-2	Sequence 2, Appli
41	420.5	13.2	436	4	US-09-523-849-2	Sequence 2, Appli
42	417.5	13.1	433	2	US-08-913-050A-7	Sequence 7, Appli
43	417.5	13.1	433	2	US-08-749-902-5	Sequence 5, Appli
44	413	13.0	1050	4	US-09-428-711A-16	Sequence 16, Appl
45	409.5	12.8	456	1	US-08-464-164-2	Sequence 2, Appli

ALIGNMENTS

RESULT 1
US-09-930-181-2
; Sequence 2, Application US/09930181
; Patent No. 6455292
; GENERAL INFORMATION:
; . APPLICANT: Origene Technologies
; . TITLE OF INVENTION: Full-Length Serine Protein Kinase in Brain and Pancreas
; . FILE REFERENCE: 16U 101 V1
; . CURRENT APPLICATION NUMBER: US/09/930.181
; . CURRENT FILING DATE: 2001-08-16
; . NUMBER OF SEQ ID NOS: 18
; . SOFTWARE: PatentIn version 3.0
; . SEQ ID NO 2
; . LENGTH: 668
; . TYPE: PRT
; . ORGANISM: Homo sapiens
US-09-930-181-2

Query Match	100.0%	Score	3188	DB	4	Length	668
Best Local Similarity	100.0%	Pred. No.	6.8e-218				
Matches	608	Conservative	0	Mismatches	0	Indels	0
Gaps	0						
Qy	1	MKVERETAILKLEHPPHVLKLDHVVYENKYLVLVLEHVS	GGELFDYLVKKGR	LTPKEARK	60		
Db	61	MKVERETAILKLEHPPHVLKLDHVVYENKYLVLVLEHVS	GGELFDYLVKKGR	LTPKEARK	120		
Qy	61	FFROIISALDFCHSHSICHRLDKPENLLDEKKNIRIADFG	MASLQVDSLL	ETSCGSPH	120		
Db	121	FFROIISALDFCHSHSICHRLDKPENLLDEKKNIRIADFG	MASLQVDSLL	ETSCGSPH	180		
Qy	121	YACPEVIRGEKYDGRKADVMSCGVILFALLV	GVLPDDNLRQLLEK	YKRGVFMHPFIP	180		
Db	181	YACPEVIRGEKYDGRKADVMSCGVILFALLV	GVLPDDNLRQLLEK	YKRGVFMHPFIP	240		
Qy	181	PDCOSLARGMEVDAAARLTLEHIQKHWYIGGKNEPEPE	PIPRKVQIRSL	SLSDIDP	240		
Db	241	PDCOSLARGMEVDAAARLTLEHIQKHWYIGGKNEPEPE	PIPRKVQIRSL	SLSDIDP	300		
Qy	241	DVLDSMHSGLCFDRDNKLQDLLSEENQEKMIYFLL	LDKRYPSQ	DEDLPPRNEIDP	300		
Db	301	DVLDSMHSGLCFDRDNKLQDLLSEENQEKMIYFLL	LDKRYPSQ	DEDLPPRNEIDP	360		
Qy	301	PRKVDSPMLNRHGRKRRPERKSMVLSVTDGGS	VPARRAIEA	MAHQGRSIS	SGASSGL	360	
Db	361	PRKVDSPMLNRHGRKRRPERKSMVLSVTDGGS	VPARRAIEA	MAHQGRSIS	SGASSGL	420	
Qy	361	STPSLSSPRVTPHSPRGSLPTPKGTPVHTPKES	PAGTNP	TPSPSPSV	GVGVPWRARLN	420	
Db	421	STPSLSSPRVTPHSPRGSLPTPKGTPVHTPKES	PAGTNP	TPSPSPSV	GVGVPWRARLN	480	

QY 421 SIKNSFLGSPFRHRRKLVQPTPEEMSNLTPESSPELAKSWFGNFISLEKEEQIFVVIKD 480
DB 481 SIKNSFLGSPFRHRRKLVQPTPEEMSNLTPESSPELAKSWFGNFISLEKEEQIFVVIKD 540
QY 481 KPLSSIKADIIVHAFSLPSLSHSHSVISQTSFRAEYKATGGPAVFOKPVKQFQVDITYTGGGE 540
DB 541 KPLSSIKADIIVHAFSLPSLSHSHSVISQTSFRAEYKATGGPAVFOKPVKQFQVDITYTGGGE 600
QY 541 AQKNGIYSVFTLLSGSPRRFKRVETIQALLSTHDPDPAQHLSDTTNCMENMTGRLS 600
DB 601 AQKNGIYSVFTLLSGSPRRFKRVETIQALLSTHDPDPAQHLSDTTNCMENMTGRLS 660
QY 601 KCGIIPKS 608
DB 661 KCGIIPKS 668

RESULT 2
US-09-930-181-4
; Sequence 4, Application US/09930181
; Patent No. 6455292
; GENERAL INFORMATION:
; APPLICANT: Origene Technologies
; TITLE OF INVENTION: Full-Length Serine Protein Kinase in Brain and Pancreas
; FILE REFERENCE: 16U 101 V1
; CURRENT APPLICATION NUMBER: US/09/930,181
; CURRENT FILING DATE: 2001-08-16
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 4
; LENGTH: 585
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-930-181-4

Query Match 95.0%; Score 3028; DB 4; Length 585;
Best Local Similarity 100.0%; Pred. No. 1.2e-206;
Matches 577; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 32 YLVLEHVSQGLFDYLVKKGRLTPKEARKFFQIISALDFCHSHSICHRLDKPENLLLDE 91
DB 9 YLVLEHVSQGLFDYLVKKGRLTPKEARKFFQIISALDFCHSHSICHRLDKPENLLLDE 68
QY 92 KNNIRIADFGMASLQVGSLLTSCGSPHYACPEVIRGEKYDGRKADVMSQGVILFALLV 151
DB 69 KNNIRIADFGMASLQVGSLLTSCGSPHYACPEVIRGEKYDGRKADVMSQGVILFALLV 128
QY 152 GALPDDDDNLRLQLLEKVKRGVFMHPHFIPPCQSLLRGMI EVDAAARLTLEHIQKHIWI 211
DB 129 GALPDDDDNLRLQLLEKVKRGVFMHPHFIPPCQSLLRGMI EVDAAARLTLEHIQKHIWI 188
QY 212 GKGNEPEQPIPRKVQIRSLPSLEDIDPDLSDMSHSLGCFDRDKLQDLLSEENQEK 271
DB 189 GKGNEPEQPIPRKVQIRSLPSLEDIDPDLSDMSHSLGCFDRDKLQDLLSEENQEK 248
QY 272 MIYFLLDRKERYPSQEDIDLPRNEIDPPRKRVDSPLNRRHGRKRRPERKSMVLSVTDG 331
DB 249 MIYFLLDRKERYPSQEDIDLPRNEIDPPRKRVDSPLNRRHGRKRRPERKSMVLSVTDG 308
QY 332 GSPVPAARAIEMAHQGRSRISGASSGLSTSPSSPVTPHPSRGSPLTPKGTPTVHT 391
DB 309 GSPVPAARAIEMAHQGRSRISGASSGLSTSPSSPVTPHPSRGSPLTPKGTPTVHT 368
QY 392 PKESPAGTPNTPPSSPSVGGVPMARLNSIKNSFLGSPFRHRRKLVQPTPEEMSNLTPE 451
DB 369 PKESPAGTPNTPPSSPSVGGVPMARLNSIKNSFLGSPFRHRRKLVQPTPEEMSNLTPE 428
QY 452 SSPELAKSWFGNFISLEKEEQIFVVIKDKPLSSIKADIIVHAFSLPSLSHSHSVISQTSFR 511
DB 429 SSPELAKSWFGNFISLEKEEQIFVVIKDKPLSSIKADIIVHAFSLPSLSHSHSVISQTSFR 488
QY 512 AYEKATGGPAVFOKPVKQFQVDITYTGGGEAQKNGIYSVFTLLSGSPRRFKRVETIQOA 571

DB 489 AYEKATGGPAVFOKPVKQFQVDITYTGGGEAQKNGIYSVFTLLSGSPRRFKRVETIQOA 548
QY 572 QLLSTHDPDPAQHLSDTTNCMENMTGRLSKCGIIPKS 608
DB 549 QLLSTHDPDPAQHLSDTTNCMENMTGRLSKCGIIPKS 585

RESULT 3
US-09-930-181-17
; Sequence 17, Application US/09930181
; Patent No. 6455292
; GENERAL INFORMATION:
; APPLICANT: Origene Technologies
; TITLE OF INVENTION: Full-Length Serine Protein Kinase in Brain and Pancreas
; FILE REFERENCE: 16U 101 V1
; CURRENT APPLICATION NUMBER: US/09/930,181
; CURRENT FILING DATE: 2001-08-16
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 17
; LENGTH: 603
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-930-181-17

Query Match 94.6%; Score 3015; DB 4; Length 603;
Best Local Similarity 99.7%; Pred. No. 1.1e-205;
Matches 574; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 12 LIHHPVHLKLDVYENKKYLYLVLEHVSQGLFDYLVKKGRLTPKEARKFFQIISALDF 71
DB 1 LIHHPVHLKLDVYENKKYLYLVLEHVSQGLFDYLVKKGRLTPKEARKFFQIISALDF 60
QY 72 CHSHSICHRLDKPENLLDEKNNIRIADFGMASLQVGSLLTSCGSPHYACPEVIRGEK 131
DB 61 CHSHSICHRLDKPENLLDEKNNIRIADFGMASLQVGSLLTSCGSPHYACPEVIRGEK 120
QY 132 YDGRKADVMSQGVILFALLVQALPFDNDNLRLQLLEKVKRGVFMHPHFIPPCQSLLRGMI 191
DB 121 YDGRKADVMSQGVILFALLVQALPFDNDNLRLQLLEKVKRGVFMHPHFIPPCQSLLRGMS 180
QY 192 EVDAAARLTLEHIQKHIWIYIGGKNEPEPEQPIPRKVQIRSLPSLEDIDPDLSDMSHSLGC 251
DB 181 EVDAAARLTLEHIQKHIWIYIGGKNEPEPEQPIPRKVQIRSLPSLEDIDPDLSDMSHSLGC 240
QY 252 FRDNKLLQDLLSEENQEKMIYFLLDRKERYPSQEDIDLPRNEIDPPRKRVDSPLN 311
DB 241 FRDNKLLQDLLSEENQEKMIYFLLDRKERYPSQEDIDLPRNEIDPPRKRVDSPLN 300
QY 312 RHGKRPRERKSMVLSVTDGSPVPARRAIEMAHQGRSRISGASSGLSTSPSSPVTPHPSRGS 371
DB 301 RHGKRPRERKSMVLSVTDGSPVPARRAIEMAHQGRSRISGASSGLSTSPSSPVTPHPSRGS 360
QY 372 PHSPRGSPLTPKGTPTVHTPKESPAGTPNTPPSSPSVGGVPMARLNSIKNSFLGSPR 431
DB 361 PHSPRGSPLTPKGTPTVHTPKESPAGTPNTPPSSPSVGGVPMARLNSIKNSFLGSPR 420
QY 432 FHRKLVQPTPEEMSNLTPESSPELAKSWFGNFISLEKEEQIFVVIKDKPLSSIKADIIV 491
DB 421 FHRKLVQPTPEEMSNLTPESSPELAKSWFGNFISLEKEEQIFVVIKDKPLSSIKADIIV 480
QY 492 HAFSLPSLSHSHSVISQTSFRAEYKATGGPAVFOKPVKQFQVDITYTGGGEAQKNGIYSVT 551
DB 481 HAFSLPSLSHSHSVISQTSFRAEYKATGGPAVFOKPVKQFQVDITYTGGGEAQKNGIYSVT 540
QY 552 FTLLSGSPRRFKRVETIQALLSTHDPDPAQHLSD 587
DB 541 FTLLSGSPRRFKRVETIQALLSTHDPDPAQHLSD 576

RESULT 4
US-09-523-849-36
; Sequence 36, Application US/09523849

Patent No. 6458561
 GENERAL INFORMATION:
 APPLICANT: Bandman, Olga
 APPLICANT: Molteni, Angela
 APPLICANT: Magnaghi, Paola
 APPLICANT: Bogotti, Roberta
 APPLICANT: Scacheri, Emanuela
 APPLICANT: Isacchi, Antonella
 APPLICANT: Hodgson, Dave
 TITLE OF INVENTION: HUMAN NIM1 KINASE
 FILE REFERENCE: PC-0009 US
 CURRENT APPLICATION NUMBER: US/09/523,849
 CURRENT FILING DATE: 2000-03-13
 NUMBER OF SEQ ID NOS: 39
 SOFTWARE: PERL Program
 SEQ ID NO 36
 LENGTH: 745
 TYPE: PRT
 ORGANISM: Homo sapiens
 NAME/KEY: misc feature
 FEATURE:
 OTHER INFORMATION: GenBank Accession No. 6458561 g1749794
 US-09-523-849-36

Query Match 21.6%; Score 688.5; DB 4; Length 745;
 Best Local Similarity 28.1%; Pred. No. 7.8e-41;
 Matches 202; Conservative 108; Mismatches 225; Indels 183; Gaps 21;

QY 2 KVERETAILKLEHPHVLKLDHVDYENKKYLVLVLEHVSQGELEFDYLVKKGLTPEAKRF 61
 DB 63 KLFREVRIMKVLNHPNIVKLFVETETKLYLVMEYASGGEVFDYLVAGHMKKEKARAK 122
 QY 62 FROIISALDFCHSHSICHRLDKPENLLLDKKNIRIADFGMASLQVGSLLLETSCGSPHY 121
 DB 123 FROIIVSAVQYCHQKFIHVDLKAENLLLDADNMNIIADFGFSNEFTGKLDTCGSPY 182
 QY 122 ACEVIRGKYDGRKADVWSCGVILFALLVGLPDDNLRLQLEKKYKRGVFMHPHFIPP 181
 DB 183 AAPELFQGGKYDGEVDVWVSLGVILYTLVSGSLPFDGQNLKELRERVLRGKYRIPFYVMT 242
 QY 182 DCQSLRGMIEVDAAARLLTLEHIOKHIWYIGGKNEPEEQIPKQVIRSLPSLEDIDPD 241
 DB 243 DCENLLKKFLINPNSKRGTEQIMKDRMNMNVGHEDDE-----LKPVEPLPDYK--DPR 294
 QY 242 VLDSMHSGLCFDRNKLQDLSEENOEKMIYFLLDRKERYPSQEDDLPPRNEID-- 299
 DB 295 RTLEWVSMGYTREE---IQDSLQVQRYNEVMATYLLGYKSELEGDTITLKPRPSADLT 351
 QY 300 -----PPRKRVDS-----PMLNRHGK-----RRPE-----R 320
 DB 352 NSSAQPSPHKVQSVSANPKQRRFSDQAGPAIPTSNYSKKTQSNNAENKRPEDRESGR 411
 QY 321 KSM-----EVLVTDG---GSPVARRAIEWA---OHQORS 350
 DB 412 KASSTAKVPASPLGLERKKTTPSTNSVLSTSTNRNSPLERASLQASIQNGKDS 471
 QY 351 RSTSG--ASSGLSTPLSPRPVTPH-----PSPRGSPPLTPKGTVPVHT 391
 DB 472 LTPWGSASTASAAVSAARPRQKMSASVHPNKASGLPPTESNCEVPRPSTAQRV 531
 QY 392 PKESPA-----GTPNPT-----PPSPSPSVGV 413
 DB 532 PVASPSAHNIISSGGAPDRTNFPRGVSSRSTFHAGQLQVRDQNLQVGVTPASPSGHSQ 591
 QY 414 PWRARLNSIKNSFLGSPRRHRLQVPTPEMSNLTP-----ESSPEL 456
 DB 592 GRGASGSIIFSFP--TSKFVRRLNLEPESKDRVETLRPHVVGSGNDKEEPEAKPRS 649
 QY 457 AKKSW--FGNFISLEKEEQIFVVIKDKPLSSIKADIVHAFPLSLHSVISQTSFRAEYK 515
 DB 650 LRTWNSKNTSSMEPNEMWREIRKVLVDANSQCSSELHEKYMLL--CMHTPGHEDF----- 702
 QY 516 ATGGPAVQKPKVFQVDITYTEGGEAQKENGIVSYVTFTLLSGPSRRFKRVVETIIQAOL 573

US-09-984-890-2
 RESULT 5
 US-09-984-890-2
 Sequence 2, Application US/09984890
 Patent No. 6492156
 GENERAL INFORMATION:
 APPLICANT: VAN, Chunhua et al.
 TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC
 TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES
 TITLE OF INVENTION: THEREOF
 FILE REFERENCE: CL001306
 CURRENT APPLICATION NUMBER: US/09/984,890
 CURRENT FILING DATE: 2001-10-31
 NUMBER OF SEQ ID NOS: 4
 SOFTWARE: FastSeq for Windows Version 4.0
 SEQ ID NO 2
 LENGTH: 724
 TYPE: PRT
 ORGANISM: Homo sapiens
 US-09-984-890-2

Query Match 21.5%; Score 685.5; DB 4; Length 724;
 Best Local Similarity 28.9%; Pred. No. 1.2e-40;
 Matches 194; Conservative 103; Mismatches 230; Indels 145; Gaps 19;

QY 2 KVERETAILKLEHPHVLKLDHVDYENKKYLVLVLEHVSQGELEFDYLVKKGLTPEAKRF 61
 DB 96 KLFREVRIMKVLNHPNIVKLFVETETKLYLVMEYASGGEVFDYLVAGHMKKEKARAK 155
 QY 62 FROIISALDFCHSHSICHRLDKPENLLLDKKNIRIADFGMASLQVGSLLLETSCGSPHY 121
 DB 156 FROIIVSAVQYCHQKFIHVDLKAENLLLDADNMNIIADFGFSNEFTGKLDTCGSPY 215
 QY 122 ACEVIRGKYDGRKADVWSCGVILFALLVGLPDDNLRLQLEKKYKRGVFMHPHFIPP 181
 DB 216 AAPELFQGGKYDGEVDVWVSLGVILYTLVSGSLPFDGQNLKELRERVLRGKYRIPFYVMT 275
 QY 182 DCQSLRGMIEVDAAARLLTLEHIOKHIWYIGGKNEPEEQIPKQVIRSLPSLEDIDPD 241
 DB 276 DCENLLKKFLINPNSKRGTEQIMKDRMNMNVGHEDDE-----LKPVEPLPDYK--DPR 327
 QY 242 VLDSMHSGLCFDRNKLQDLSEENOEKMIYFLLDRKERYPSQEDDLPPRNEIDDP 301
 DB 328 RTLEWVSMGYTREE---IQDSLQVQRYNEVMATYLLGYKSELEGDTITLKPRPSADLT 384
 QY 302 RKRVDSP--MLNRHGKRPERSMEVLSVTDGSPVPA-----RRATEMA 344
 DB 385 NSSAQPSPHKVQSVSANPKQRRFS---DQAGPAIPTSNYSKKTQSNNAENKRPEDR 440
 QY 345 QHGQSRISIGSGLSTSP---LSSPRVTPHS-----PRGSP-- 381
 DB 441 ESGRKASS---TAKVPASPLGLERKKTTPSTNSVLSTSTNRNSPLERASLQAS 496
 QY 382 -----PTPKGTVPVHTPK-----ESPA GTNPT----- 403
 DB 497 STQNGKDTAQRVDPVAFSPSAHNIISSGGAPDRTNFPRGVSSRSTFHAGQLQVRDQNL 556
 QY 404 -----PPSPSPSVGVPMARLNSIKNSFLGSPRRHRLQVPTPEMSNLTP----- 450
 DB 557 PYGVTPASPSGHSQGRRCASGSIIFSFP--TSKFVRRLNLEPESKDRVETLRPHVVGSGN 614
 QY 451 -----ESSPELAKSW--FGNFISLEKEEQIFVVIKDKPLSSIKADIVHAFPLSLHS 501
 DB 615 DKEKEEPEAKPRSIRFTWNSKNTSSMEPNEMWREIRKVLVDANSQCSSELHEKYMLL--CM 672
 QY 502 HSVISQTSFRAEYKATGGPAVQKPKVFQVDITYTEGGEAQKENGIVSYVTFTLLSGPSRR 561
 DB 673 HGTGPHEDF-----VQWMEVC-----KLPRLSLNGVRFKRVISGTSMA 710
 QY 562 FKRVVETIIQAOL 573

Db 711 FKNTASKIANEL 722

RESULT 6

US-09-984-890-4

; Sequence 4, Application US/09984890

; Patent No. 6492156

; GENERAL INFORMATION:

; APPLICANT: YAN, Chunhua et al.

; TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC

; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES

; FILE REFERENCE: CL001306

; CURRENT APPLICATION NUMBER: US/09/984.890

; CURRENT FILING DATE: 2001-10-31

; NUMBER OF SEQ ID NOS: 4

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 4

; LENGTH: 722

; TYPE: PRT

; ORGANISM: Rattus norvegicus

US-09-984-890-4

Query Match 21.5%; Score 684.5; DB 4; Length 722;

Best Local Similarity 29.3%; Pred. No. 1.4e-40;

Mismatches 195; Conservative 103; Mismatches 233; Indels 135; Gaps 18;

Qy 2 KVERETAILKLIHPHVLKLDHYVENKYLIVLHVHSGGLFDYLVKKGLTPEKARKF 61

Db 96 KLFREVRIMKVLNHPNIVKLFVETETKTLVLMVEYASGGEVFDYLVAGHGMKEKARAK 155

Qy 62 FROIISALDFCHSHSICHRDLKPELNLLDEKNNIRIADFGMASIQVGDLSLETSCGSPHY 121

Db 156 FROIIVSAVQYCHKFIHVRDLKAENLLDADNMKIADFGSNEFTGKNDLTCGSPPY 215

Qy 122 ACPEVIRGEYDKGRKADVWCGVILFALLVGLPFDNDNLKQLLEKVKRGVFMHPHIPP 181

Db 216 AAPFLQGGKYDGPEDVWLSGLVILTVSGSLPFDGQNLKELRERVLRGKRIPIFYMT 275

Qy 182 DCQSLLRGMIEVDAAARLTLEHIQKHIWYIGKNEPEPEQIPRKVQIRSLPSLEDIDPD 241

Db 276 DCENLLKFLILNPSKRGTELEQIMKDRVMNVGHEDDE-----LKPVEPLPDYK--DPR 327

Qy 242 VLDSMHSIGCFDRDNKLLQDLSSEENQEKMIYFLLDLRKERYPSQDEDLPPRNEID-- 299

Db 328 RTELMSVNGYTREE---IQDSLVGQRYNEVWATYLLGLYKSELEGDTITLKPRPSADLT 384

Qy 300 -----PPRKRVDS-----PMLNRHGK-----RRPE-----RKS 322

Db 385 NSSAPSPSHKVQSVSANPKQRSSDQAVPAIPTSNYSYKKTQSNNAENKRPBEETGRKA 444

Qy 323 MEVLVTDGSGVPARRAIEWAHQHQRSSISGASSGLSTSP----- 365

Db 445 SSTAQVP--ASPLGLDRKKTTPTPTSTNSVLSTSTNRSRNSPLDRASLQASIQNGKDS 502

Qy 366 SSPRYTPHPSPRGSLPTPKGTPVHTPKESPAGT-----PNPTTP 405

Db 503 TAPQVVPVSPAHNISSSGAPDRT--NFRKGVSSRTFFHAGQLRQVRDQNLPGVT 560

Qy 406 SSPSVGGVFWARLANSIKNSFLGSPRFRHKKLQVP--TPEMSNLTP----- 450

Db 561 ASPSGHSGQRRGASGSIKSF--TSKFVRNLNPEPSKDRVETLPHVVGGGTDKEKEE 618

Qy 451 ---ESSPELAKGW--GNFISLEKEEQIFVVIKDKPLSSIKADIVAFISLISLSHVSISQ 507

Db 619 FREAKPRSLRFTWSMKTTSSEPMEMREIRKVLDSANCSQSELHERYMLL--CVHGTGPH 676

Qy 508 TSFRAEYATGPAVQKPKVQVDITYTEGGEAQKENGIIYVTTLLSGSPRRPKRVVE 567

Db 677 ENF-----VQWEMEVC-----KLPRLSLNGVRFRKISGTSMAFNATAS 714

Qy 568 TIQAQL 573

Db 715 KIANEL 720

RESULT 7

US-08-817-832B-32

; Sequence 32, Application US/08817832B

; Patent No. 6579691

; GENERAL INFORMATION:

; APPLICANT: MANDELKOW, Eckhard, et al.

; TITLE OF INVENTION: No. 6579691el Protein Kinase (NPK-110)

; NUMBER OF SEQUENCES: 32

; CORRESPONDENCE ADDRESS:

; ADDRESSES: Marshall, O'Toole, Gerstein, Murray & Borun

; STREET: 233 S. Wacker Drive, 6300 Sears Tower

; CITY: Chicago

; STATE: Illinois

; COUNTRY: US

; ZIP: 60606-6402

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: Patent in Release #1.0, Version #1.30 (EPO)

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/817,832B

; FILING DATE: 28-APR-1997

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: WO PCT/EP95/04258

; FILING DATE: 30-OCT-1995

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: EP 94 11 7122.5

; FILING DATE: 28-OCT-1994

; INFORMATION FOR SEQ ID NO: 32:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 722 amino acids

; TYPE: amino acid

; STRANDEDNESS: single

; TOPOLOGY: linear

; MOLECULE TYPE: protein

US-08-817-832B-32

Query Match 21.4%; Score 682.5; DB 4; Length 722;

Best Local Similarity 29.3%; Pred. No. 2e-40;

Mismatches 195; Conservative 103; Mismatches 233; Indels 135; Gaps 18;

Qy 2 KVERETAILKLIHPHVLKLDHYVENKYLIVLHVHSGGLFDYLVKKGLTPEKARKF 61

Db 96 KLFREVRIMKVLNHPNIVKLFVETETKTLVLMVEYASGGEVFDYLVAGHGMKEKARAK 155

Qy 62 FROIISALDFCHSHSICHRDLKPELNLLDEKNNIRIADFGMASIQVGDLSLETSCGSPHY 121

Db 156 FROIIVSAVQYCHKFIHVRDLKAENLLDADNMKIADFGSNEFTGKNDLTCGSPPY 215

Qy 122 ACPEVIRGEYDKGRKADVWCGVILFALLVGLPFDNDNLKQLLEKVKRGVFMHPHIPP 181

Db 216 AAPFLQGGKYDGPEDVWLSGLVILTVSGSLPFDGQNLKELRERVLRGKRIPIFYMT 275

Qy 182 DCQSLLRGMIEVDAAARLTLEHIQKHIWYIGKNEPEPEQIPRKVQIRSLPSLEDIDPD 241

Db 276 DCENLLKFLILNPSKRGTELEQIMKDRVMNVGHEDDE-----LKPVEPLPDYK--DPR 327

Qy 242 VLDSMHSIGCFDRDNKLLQDLSSEENQEKMIYFLLDLRKERYPSQDEDLPPRNEID-- 299

Db 328 RTELMSVNGYTREE---IQDSLVGQRYNEVWATYLLGLYKSELEGDTITLKPRPSADLT 384

Qy 300 -----PPRKRVDS-----PMLNRHGK-----RRPE-----RKS 322

Db 385 NSSAPSPSHKVQSVSANPKQRSSDQAVPAIPTSNYSYKKTQSNNAENKRPBEETGRKA 444

Qy 323 MEVLVTDGSGVPARRAIEWAHQHQRSSISGASSGLSTSP----- 365

Db 445 SSTAQVP--ASPLGLDRKKTTPTPTSTNSVLSTSTNRSRNSPLDRASLQASIQNGKDS 502

Db 363 RQPOLRNDLSSLEVPQEIPLPCOPFRPSLLCPOQALAQVLAQIDCDLHSLQLPLFP 422
Qy 319 -----ERKNEVLVSVDGSPVARRAEMAQHQSRISGASSGLST----- 362
Db 423 LDTNCSGVFRHRSISPSLLDTAISEARQGPSLEEEOEVOEPLPGSTGRHHTLAEVSTH 482
Qy 363 - SPLSSPRVTFPHSPRGSPLTPKGT-----VHTPKESPAG-----TPNPTPPSPSV 410
Db 483 FSPINLPCIIIVSSAAVSP---SGTSSDSCLPFSASEGAGLGGGLATPGLLTSSP-- 537
Qy 411 GGVPRARLINSIKNSFLGSPFRHRRKLVQVTPPEMSN-----LTPESPPE----- 455
Db 538 -----VRLAS-----PFLGS-----QSATPVLSQAGLGLATVLPVSPQEGRRASDTS 580
Qy 456 -----LAKSNFNGFISLEKEEQIFVWIKDKPLSLSIKADIVHAFLSIPSLSHSV 504
Db 581 LTQGLKAFRQQLRNARTKGLGLNK-----TKGLARQV 614
Qy 505 ISQTSFRAEYKATGGPAVFQKP 526
Db 615 -CQSSIRG---SRGGMSTFHTP 632

RESULT 14

US-08-557-006C-40
; Sequence 40, Application US/08557006C
; Patent No. 6258547

GENERAL INFORMATION:

; APPLICANT: Beri, Rajindar K.
; APPLICANT: Carling, David
; APPLICANT: Forster, Robert A.
; TITLE OF INVENTION: NUCLEIC ACID ENCODING AMP-ACTIVATED PROTEIN KINASE
; FILE REFERENCE: NGAP/PHM37588/UST
; CURRENT APPLICATION NUMBER: US/08/557,006C
; CURRENT FILING DATE: 1996-03-05
; PRIOR APPLICATION NUMBER: PCT/GB94/01093
; PRIOR FILING DATE: 1994-05-20
; PRIOR APPLICATION NUMBER: GB 9310489.1
; PRIOR FILING DATE: 1993-05-21
; PRIOR APPLICATION NUMBER: GB 9318010.7
; PRIOR FILING DATE: 1993-08-31
; NUMBER OF SEQ ID NOS: 44
; SOFTWARE: Patent in Ver. 2.1
; SEQ ID NO 40
; LENGTH: 552
; TYPE: PRT
; ORGANISM: Rat
; FEATURE:
; NAME/KEY: Gene
; LOCATION: (1)..(1747)
; OTHER INFORMATION: Full length cDNA sequence fragment of Human AMPK -
; OTHER INFORMATION: fragment begins at nucleotide 24 and ends with
; OTHER INFORMATION: nucleotide 1765
US-08-557-006C-40

Query Match

Best Local Similarity 19.9%; Score 636; DB 3; Length 552;
Matches 174; Conservative 95; Mismatches 202; Indels 142; Gaps 21;

Qy 2 KVERIATIKLEHPHVHLKHVDYENKKYLVYLVHVSGLGELFDYLVKKGRITPKAEKRF 61
Db 60 KIKRBIQNLKPRPHIHKLYQVISTPTDFVMBYVSGGELFDYICKHGRVEVEARL 119
Qy 62 FQIISALDFCHSHICHRLDKPENLLDEKNIRIADFGMASLQVDSGLSLETSCGSPHY 121
Db 120 FQIILSANDYCHRHVVRHDLKPNVLLDAQNAXIADFGLSNMMSDGEFLRTSCGSPNY 179
Qy 122 ACPEVIREKYDGRKADVWSCVILFALLVGPALPDNDNLRLLEKVKRGVPHPHFTPP 181
Db 180 ARAPEVISGRLYAGPEVDIWSGVILYALLCGTLPEDDEHVPFLFKIRGGVFYIPEYLR 239
Qy 182 DCQSLRGVIEVDAARRLTLEHQHIVYIGGKNEPEEQIPRKVQIRSLPSLEDIDPD 241

Db 240 STATLMMHMLQVDPLKRATIKDIREHEWF-----KQDLPSYL-----FPEDPSYDAN 286
Qy 242 VLDS-----MHSGLCFRDNKLLQDILLS-EENOEKMIYELLDRKERYPSQEDS---- 290
Db 287 VIDDEAVKVCSEKFEK--TESEVMNSLYSGDQDQAVAYHLIIDNR-RIMQAASEFYLA 343
Qy 291 DLPPRNEIDPPRRKRVDSPLMNRHGKRPERKMEVLVSVDGSPVARRAEMAQHQGRS 350
Db 344 SSPPTG-----SFMD-----DNAMEH----- 358
Qy 351 RSISGASSGLSTSPSSPRVTFPHSPRGSPL--PPPKGTGVHTPKESPAGTNPPTPSP 408
Db 359 -----IPPLGLKPHPE-RMPPLIADSPK-----RCPLDALNTTKPKSL 395
Qy 409 SVGVGVPWRARLINSIKNSFLGSPFRHRRKLVQVTPPEMSN-----TPESPPELAKSKWFGN 464
Db 396 AVKAKWHGIRSQSKPYDIMAETVRAMKQDFEWKVVNAYHLRVRKKNPVT-----GN 449
Qy 465 FISLEKEEQIFVWIKDKPLSLSIKADIVHAFLSIPSLSHSVISQTSFRAEYKATGGPAVFQ 524
Db 450 YVQMSL--QLYLV-----DNRSYLLDFKSIDDEVQORSSTPQRSCTSAAGLH 496
Qy 525 KPVKFOVDITYTEGGAQKENGIIYVTFLLSGPSRRFRKVVETIOAQLLSTHDPAAQH 584
Db 497 RP-RSSVDSSTAENHS-----LSQ-----SLTGLTGLTSLSSASPLGSH 535
Qy 585 LSDTTN-CMEMMT 596
Db 536 TMDFFENCASLIIT 548

RESULT 15

US-09-101-146-1
; Sequence 1, Application US/09101146
; Patent No. 6124125

GENERAL INFORMATION:

; APPLICANT: Dartmouth College, St. Vincents Institute of
; APPLICANT: Medical Research, Kemp et al.
; TITLE OF INVENTION: No. 6124125el AMP Activated Protein Kinase
; NUMBER OF SEQUENCES: 64
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Jane Massey Licata, Esq.
; STREET: 66 E. Main Street
; CITY: Marlton
; STATE: NJ
; COUNTRY: USA
; ZIP: 08053

COMPUTER READABLE FORM:

; MEDIUM TYPE: DISKETTE, 3.5 INCH, 1.44 Mb STORAGE

; COMPUTER: IBM PC

; OPERATING SYSTEM: WINDOWS 95

; SOFTWARE: WORDPERFECT 6.0 FOR WINDOWS

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/09/101,146

; FILING DATE: October 7, 1998

; CLASSIFICATION: 435

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: PN7450

; FILING DATE: 8 JAN 1996

; ATTORNEY/AGENT INFORMATION:

; NAME: Jane Massey Licata

; REGISTRATION NUMBER: 32,257

; REFERENCE/DOCKET NUMBER: DC-0050

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (856) 810-1515

; TELEFAX: (856) 810-1454

; INFORMATION FOR SEQ ID NO: 1:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 345

; TYPE: Amino acid

; TOPOLOGY: Linear

US-09-101-146-1

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 26, 2003, 11:57:16 ; Search time 16.6771 Seconds
(without alignments)
3506.035 Million cell updates/sec

Title: US-10-054-579-4
Perfect score: 3188
Sequence: 1 MKVERETAILKLEHPHVLK.....TNCWEMMTGRSLKXGIIPKS 608

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283308 seqs, 96168682 residues

Total number of hits satisfying chosen parameters: 283308

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

1: Pir1.*
2: Pir2.*
3: Pir3.*
4: Pir4.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES:

Result No.	Score	Query Match	Length	DB ID	Description
1	1536	48.2	887	2 T20941	hypothetical prote
2	722	22.6	1518	2 S37928	probable purine nu
3	698	21.9	401	2 B30120	SNF1-related prote
4	698	21.9	512	1 J21446	serine/threonine-s
5	687.5	21.6	745	2 G01025	serine/threonine p
6	686	21.5	504	2 T10449	probable serine/th
7	685.5	21.5	775	2 T38929	changed division r
8	681.5	21.4	512	2 T52633	serine/threonine-s
9	672.5	21.1	473	1 S59941	serine/threonine-s
10	668.5	21.0	1142	2 S59359	GIN4 protein - yea
11	668	21.0	511	1 A56009	serine/threonine-s
12	662.5	20.8	576	2 T41587	probable carbon ca
13	661.5	20.7	774	2 T48609	probable serine/th
14	660.5	20.7	562	2 T29858	hypothetical prote
15	655	20.5	915	2 S74283	probable protein k
16	654.5	20.5	713	2 S27966	probable serine/th
17	650	20.4	512	2 T07788	probable serine/th
18	649	20.4	798	2 J27500	gk protein - chic
19	642.5	20.2	504	2 T07415	probable serine/th
20	636	19.9	552	1 A53621	(hydroxymethylglut
21	632	19.8	552	1 S51025	probable serine/th
22	625	19.6	1192	2 T18611	protein H3923.1 [
23	624	19.6	1246	2 G99287	SNF-related kinase
24	620	19.4	472	2 B30100	hypothetical prote
25	616.5	19.3	1558	2 T29253	serine/threonine-s
26	616	19.3	633	1 A26030	serine/threonine-s
27	611.5	19.2	502	1 A41361	serine/threonine-s
28	609	19.1	602	2 S72513	FOG2 protein - yea
29	606	19.0	513	1 S60303	serine/threonine-s

ALIGNMENTS

RESULT 1

T20941

hypothetical protein F15A2.6 - Caenorhabditis elegans

C;Species: Caenorhabditis elegans

C;Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 18-Feb-2000

C;Accession: T20941

R;Gregory, J.

submitted to the EMBL Data Library, March 1996

A;Reference number: Z19349

A;Accession: T20941

A;Status: preliminary; translated from GB/EMBL/DDBJ

A;Molecule type: DNA

A;Residues: 1-887 <WIL>

A;Cross-references: EMBL:Z70207; PIDN:CAA94127.1; GSPDB:GN00028; CESP:F15A2.6

A;Experimental source: clone F15A2

C;Genetics:

A;Gene: CESP:F15A2.6

A;Map position: X

A;Introns: 32/1; 63/3; 92/2; 139/2; 189/3; 328/2; 448/2; 516/3; 604/2; 684/3; 735/2; 777/

Query Match 48.2%; Score 1536; DB 2; Length 887;

Best Local Similarity 49.0%; Pred. No. 5.6e-56;

Matches 340; Conservative 71; Mismatches 133; Indels 150; Gaps 17;

Qy 2 KVERETAILKLEHPHVLKLDHYVYENKYLVLVLEHVSGLGELFDYLVKGRGLTPKGRKF 61

Db 63 KVERETAILKLEHPHVLKLDHYVYENKYLVLVLEHVSGLGELFDYLVKGRGLTPKGRKF 122

Qy 62 FROIISALDFCHSHSCHRDLKPENLLLDKNNIRIADFGMASLQVDSLLETSCGSPHY 121

Db 123 FROIISALDFCHAHNCHRDLKPENLLLDKNNIRIADFGMASLQVDSLLETSCGSPHY 182

Qy 122 ACEFVIRGEKYGCKKADVMSCGVILFALLVGLPFDNNLRQLLEKVKRGVFEMHPFP 181

Db 183 ACEFVIRGEKYGCKKADVMSCGVILFALLVGLPFDNNLRQLLEKVKRGVFEMHPFP 242

Qy 182 DCQSLLRGMTEVDAARLTLHTIQRKIWIYG-KGNEPEPEQPIPRKQVIRSLPSEDIDP 240

Db 243 DVQSLLRGMTEVDPGRKSLADYVFKHPWVGTYTKADPELPLMSQVQVTHVIFGEDSIDP 302

Qy 241 DVLDSSHSLGCFDRNKLQDLISEENQEMKLYFLLDRKERYPSQDEDLDPNRI-- 298

Db 303 DVLHNMNLCGCFCKQKLNELLSPKHTKMYFYLLDRKRPPAQEDD-----TEIVL 357

Qy 299 -----DPRKKEVDSPMLNR-----HGKRR-----PE 319

Db 358 RGAQNNDPPKKTDSRTSRYPMGSIADGSPINPRKTYGNQKSGHSLGSGSPSTESPR 417

Qy 320 RKMEVLSVTDGGS-----PVPARRAIEAHO----- 346

Db 418 SSTEDLFGSSSSSYNARAGEDRDRGRSASRSNTHYVYTPQVDPQTLAAARHVRDAQE 477

protein kinase - m

hypothetical prote

serine/threonine p

serine/threonine-s

protein kinase kin

protein kinase 1 -

probable serine/th

p698g3 protein - A

protein kinase AK2

probable protein k

hypothetical prote

serine/threonine-s

probable serine/th

hypothetical prote

probable serine/th

serine/threonine-s

```
QY 347 -----GORSRSICASSGLSTPLSSPRVTHPSGRSPLTPKGTPTVHTP 392
Db 478 RRESRDSGRGSRKESKDKSDKSSCKNDASSSTSSVPH-----KYSPPSVN 526
QY 393 KES---PACTPNTPTPSS-----PSVGGV--PWRLNSIKNSFLGSPRFRHRK 438
Db 527 SESVSVSSSTMNSTNGSTNSLIAGNSQTSIGSTSGPWSKLNINKNSFLGTGRFRHRK-- 584
QY 439 VTPPEMSNLTPTPSSPB-----LAKSWFCNF---ISLEKEQIFVWIKDPLSSI 486
Db 585 -----MSGTAESDESDQMDITDLVKKSWFGSLASMSVERDTHCVPVQGTLSI 638
QY 487 KADIVHAFISPLSHSVISQTSFRAYK--ATGGPAVFOKPVKFOVDITYTE-----G 539
Db 639 KAEILRAFLQIHELHSHVQCNCFRVEYKGPVGVSGVSRGKMNVDIIPSPQVVIAG 698
QY 540 EAKENGIVSVTFTLLSGPSRRFRVVTIOAOL 573
Db 699 ETP---TYVVQFVLLAGPVRRFKRLVEHLSAIL 728

RESULT 2
S37928
probable purine nucleotide-binding protein YKL101w - yeast (Saccharomyces cerevisiae)
N;Alternate names: protein YKL453
C;Species: Saccharomyces cerevisiae
C;Date: 03-May-1994 #sequence_revision 03-May-1994 #text_change 24-Sep-1999
C;Accession: S37928; S39084
R;Cheret, G.; Fukuhara, H.; Bolotin-Fukuhara, M.; Daignan-Fornier, B.; Pallier, C.; Puzos, V.;
submitted to the Protein Sequence Database, March 1994
A;Reference number: S37920
A;Accession: S37928
A;Molecule type: DNA
A;Residues: 1-1518 <CHR>
A;Cross-references: EMBL:Z28101; NID:g486168; PIDN:CAA81941.1; PID:g486169; MIPS:YKL101w
A;Experimental source: strain S288C
R;Pallier, C.; Valens, M.; Puzos, V.; Fukuhara, H.; Cheret, G.; Sor, F.; Bolotin-Fukuhara,
Yeast 9, 1149-1155, 1993
A;Title: DNA sequence analysis of a 17 kb fragment of yeast chromosome XI physically located
protein kinases.
A;Reference number: S39084; MUID:94078677; PMID:8256524
A;Accession: S39084
A;Status: translation not shown
A;Molecule type: DNA
A;Residues: 1-1518 <PAL>
A;Cross-references: EMBL:X71133; NID:g431205; PIDN:CAA50456.1; PID:g431215
A;Experimental source: strain S288C
C;Genetics:
A;Gene: SGD:HSL1
A;Cross-references: SGD:S0001584; MIPS:YKL101w
A;Map position: 11L
C;Superfamily: unassigned Ser/Thr or Tyr-specific protein kinases; protein kinase homolog
F;Keywords: ATP; P-loop; purine nucleotide binding; serine/threonine-specific protein kinase
F;79-369/Domain: protein kinase homolog <KIN>
F;79-86/Region: nucleotide-binding motif A (P-loop)
F;87-95/Region: protein kinase ATP-binding motif
F;85/Binding site: ATP/GTP (Lys) #status predicted

Query Match 22.6%; Score 722; DB 2; Length 1518;
Best Local Similarity 30.4%; Pred. No. 1.6e-22;
Matches 216; Conservative 98; Mismatches 250; Indels 146; Gaps 19;

QY 3 VERIATILKLEHPLVHLKLDVYENKYLVLVLEHVSGLGELFDYLVKKGRITPKAEAKFF 62
Db 161 IERIVIVKLSHTNVNMFALFEWENKSELVLVLEVDGELFDYLVKGLKLPREATHYF 220
QY 63 RQIISALDFCHSHSICHRDLKPNLLDKRN-NIRIADFGMASIQVGDLSLETSCGSPHY 121
Db 221 KQIVGVSVSYCHSNICHRDLKPNLLDKRNRIKIADFGMAELPNKLLKTSCTGSPHY 280
QY 122 ACPEVIRGEKIDGRKADYVSCGVIFALLVGVLPDDNLRQLLEKVKRGVFMHPHP 181
Db 281 ASPEIVMGRPYHGGPSDWSVCGIVLFTLLTGLHPFDDNLIKLLLVKQSGKYQMPNLS 340
```

```
QY 182 DCQSILRGMIWDAARRLTLEHIOKHIWYIGKKEPEPEQPIPKVQ----- 228
Db 341 EARDLSIKLVIDPEKRITTOEILKH--PLIKKYDDLPVNVLRKRNKDNWARGKNSDL 398
QY 229 -----IRSLPSLEDIDPDVLDMSHSLGCFDRNKLQDLJSEENQEKMIYFLLLD 279
Db 399 HLLNNVSPSIVTLHSGBEISILSLQILWHGVSRELITAKLLQKPMSEKLFYSLLIQ 458
QY 280 RKERY-----PSQDEEDLPPRNEIDPPRKRVDSPMLNRHGKRRP--KSMVLSV-TDG 331
Db 459 YKORHSISLSSSENKSKATSESSVNEPRIYASKTANNTGLRSENNDVKTLHSLHSIED 518
QY 332 GSPVPARAIE-----MAHQGRSRS-----ISCASSG 359
Db 519 TSTVQNNAITGVNTEINAPVLAQKSQFSINTLSQESDKAEAEAVTLPPAIPFNASS 578
QY 360 -----LSTPSLSPRVTHPSGRSPLTPKGTPTVHTPKESPAAGTP 400
Db 579 RIFRNSYTSISSRRSLNSRLSASTSRETVDHONEMPLQ-LPKSPSRYSLSRA 637
QY 401 NPTPPSSPSVGVPPWRARLNSIKNSFLGSPRFRKLVQVTPPEMSNLTPESSPLAKKS 460
Db 638 IHASPTKSIH-----KSLRKNIAATVAARTLQ-----NSASKRSLYSLSQISKRS 685
QY 461 WFGNFISLEKEQIFVWIKDKPLSIKADIVHAFISPLSHSVISQTSFRA----- 512
Db 686 -----LNLND-----LLVFDLPLSKKPAENVNKSEP--HSLSDSDFEILCQILFG 732
QY 513 -----EYKATGGPAVQKPVQVDITVTEGGEAKEN-----GIYSVTPT 553
Db 733 NALDRILEEEDNEKERDQQRQNDTKSSADTFTISGVSTNKNEGEPEYPTKIEKNQFN 792
QY 554 LLSGPSRRFRKV-----VETIOAQLLSTHDPAAQHLSDTTNCMEMM 595
Db 793 MSYKSPSENNSGSLSPFPEKENTLSSYLEEQPKRAA-LSDITNSFNKN 841

RESULT 3
B90120
SNF1-related protein kinase [imported] - Guillardia theta nucleomorph
C;Species: nucleomorph Guillardia theta
A;Note: a nucleomorph is the vestigial nucleus of a eukaryotic endosymbiont
C;Date: 10-May-2001 #sequence_revision 10-May-2001 #text_change 24-May-2001
C;Accession: B90120
R;Douglas, S.; Zauner, S.; Fraunholz, M.; Beaton, M.; Penny, S.; Deng, L.T.; Wu, X.; Reit
Nature 410, 1091-1096, 2001
A;Title: The highly reduced genome of an enslaved algal nucleus.
A;Reference number: A99082; MUID:11323671; PMID:11323671
A;Accession: B90120
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-401 <DOU>
A;Cross-references: GB:AF080301; NID:g13794278; PIDN:AAK39655.1; GSPDB:GN00152
C;Genetics:
A;Gene: kin(gmf2)
A;Map position: 3
A;Genome: nucleomorph
C;Keywords: nucleomorph

Query Match 21.9%; Score 698; DB 2; Length 401;
Best Local Similarity 42.7%; Pred. No. 4.1e-22;
Matches 138; Conservative 77; Mismatches 86; Indels 22; Gaps 6;

QY 1 MKVERIATILKLEHPLVHLKLDVYENKYLVLVLEHVSGLGELFDYLVKKGRITPKAEAK 60
Db 59 LKTOREISVWKMPLPHVHIKIYDVLSDSKYLFILIIYASKGELFNLYLVEKRLNREALK 118
QY 61 FFRQIISALDFCHSHSICHRDLKPNLLDKRNIRIADFGMASIQVGDLSLETSCGSPH 120
Db 119 FFHEIISGLGYCHKRICHRDLKLENILLDMKQIKIADFGMASLIPNIMLTKFCGSPH 178
QY 121 YACPEVIRGEKIDGRKADYVSCGVIFALLVGVLPDDNLRQLLEKVKRGVFMHPHP 178
```



```
Db 179 YASPEVSNPEPNGIKADWSGIIILYALVVGKLPYDEENDNRKLFNFKIRPEPRIPRI 238
Qy 179 IPPDCQSLRGMIEVDAARLLTLEHIQKHIWYIGGKNEPEPEQPIPRKVQIRSLP----- 233
Db 239 IHPDCRDLIRALLTINPKRIIDKIKHPWI---KSSPLPES---CRISQNLSPASVQ 292
Qy 234 SLEIDIDPVLDSMHSIGCFRDRNKLQDLSEENQEKMIYFLLLDRKERYPSQEDDL 292
Db 293 NPISNPDPEIISFLPLQVADKTKLGCILSTKPKPSFLVLYRQLEWRKMKMDSVRSNLI 352
Qy 293 PPRNEIDPPRKRVDSPLNRHCK 315
Db 353 ---NEANPKKK-----LFPQSK 367

RESULT 4
JC1446
serine/threonine-specific protein kinase (EC 2.7.1.1.-) AK21 - Arabidopsis thaliana
N:Alternate names: protein kinase SNF1 homolog
C:Species: Arabidopsis thaliana (mouse-ear cress)
C:Date: 30-Sep-1993 #sequence revision 30-Sep-1993 #text_change 11-Jun-1999
C:Accession: JCI1446; S58266; S66334
R:LeGuen, L.; Thomas, M.; Bianchi, M.; Halford, N.G.; Kreis, M.
Gene 120, 249-254, 1992
A:Title: Structure and expression of a gene from Arabidopsis thaliana encoding a protein
A:Reference number: JCI146; MUID:93013041; PMID:1339373
A:Accession: JCI146
A:Molecule type: DNA
A:Residues: 1-512 <LEG>
A:Cross-references: EMBL:X86966; NID:9166599; PIDN:AAA32736.1; PID:g166600
R:Thummler, F.; Kirchner, M.; Teuber, R.; Ditttrich, P.
submitted to the EMBL Data Library, May 1995
A:Description: Differential accumulation of the transcripts of 22 novel protein kinase g
A:Reference number: S58256
A:Accession: S58266
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 144-198 <THU>
A:Cross-references: EMBL:X86966; NID:928909; PIDN:CAA60529.1; PID:g928910
R:Thummler, F.; Kirchner, M.; Teuber, R.; Ditttrich, P.
Plant Mol. Biol. 29, 551-565, 1995
A:Title: Differential accumulation of the transcripts of 22 novel protein kinase genes i
A:Reference number: S66314; MUID:96123233; PMID:8534852
A:Accession: S66334
A:Molecule type: DNA
A:Residues: 144-198 <TH2>
A:Cross-references: EMBL:X86966; NID:928909; PIDN:CAA60529.1; PID:g928910
C:Comment: This enzyme plays an important role in a signal transduction cascade regulati
C:Genetics:
A:Gene: AKin10; AK21
A:introns: 64/1; 125/3; 186/3; 230/3; 292/3; 322/3; 350/3; 396/3; 475/3
C:Function:
A:Description: catalyzes the formation of peptidyl-serine-phosphate or peptidyl-threonin
C:Superfamily: AMP-activated protein kinase; protein kinase homolog
C:Keywords: ATP; magnesium; phosphotransferase; serine/threonine-specific protein kinase
F:17-27/Domain: protein kinase homolog <kin>
F:25-33/Region: protein kinase ATP-binding motif
F:48,67,142,144/Active site: Lys, Glu, Asp, Lys #status predicted
F:147,151/Binding site: magnesium (Asn, Asp) #status predicted

Query Match 21.9%; Score 698; DB 1; Length 512;
Best Local Similarity 42.7%; Pred. No. 5.2e-22;
Matches 137; Conservative 70; Mismatches 90; Indels 24; Gaps 5;

Qy 2 KVEREIALKLEIHPHVLKLDVYENKYLVLVLEHVSQGLFDYLVKKGRLTTPKEARKF 61
Db 63 KVRREIKILRFMPHPIRLVEIETPTDYLVMVEYNSGELFDYIVKGLQDEARNP 122
Qy 62 FQRIISALDFCHSHSICHRLDKPENLLDKNINRIADFGMASLQVGDLSLETSCGSPHY 121
Db 123 FQIIISGVEYCHRNWVHRDLKPNLLDSKCNVNIADFGLSNIMRDGHFLKTSQSPNY 182

A:Molecule type: DNA
A:Residues: 144-198 <TH2>
A:Cross-references: EMBL:X86966; NID:928909; PIDN:CAA60529.1; PID:g928910
C:Comment: This enzyme plays an important role in a signal transduction cascade regulati
C:Genetics:
A:Gene: AKin10; AK21
A:introns: 64/1; 125/3; 186/3; 230/3; 292/3; 322/3; 350/3; 396/3; 475/3
C:Function:
A:Description: catalyzes the formation of peptidyl-serine-phosphate or peptidyl-threonin
C:Superfamily: AMP-activated protein kinase; protein kinase homolog
C:Keywords: ATP; magnesium; phosphotransferase; serine/threonine-specific protein kinase
F:17-27/Domain: protein kinase homolog <kin>
F:25-33/Region: protein kinase ATP-binding motif
F:48,67,142,144/Active site: Lys, Glu, Asp, Lys #status predicted
F:147,151/Binding site: magnesium (Asn, Asp) #status predicted
```

```
Qy 122 ACPRVIRGEKYDGRKADVWSGVLFPALLVGCALPFDNDNLRQLLEKVKRGVFMHPPIPP 181
Db 183 AAPEVSGKLYAGPEVDVWSGVLFPALLVGCALPFDNDNLRQLLEKVKRGVFMHPPIPP 242
Qy 182 DCQSLRGMIEVDAARLLTLEHIQKHIWYIGGKNEPEPEQPIPRKVQI---RSPLPSLEDI 238
Db 243 GARDLIPRLMVLVDPKAVTIPEIKHPWF-----QAHLPYLAIVPPDPTVQQAQKI 293
Qy 239 DPVLDMSHSLGCFRDRNKLQDLSEENQEKMIYFLLLDRKERYPS-----QEDD 291
Db 294 DEELQEVINMGF---DRNHILIESLRNTQDGTVTYYLLDNRFRASSGYLGASFOETME 351
Qy 292 LPPRNEIDPPRKRVDSPLNR 312
Db 352 GTPRMH---PAESVASPVSHR 369

RESULT 5
G01025
serine/threonine protein kinase - human
C:Species: Homo sapiens (man)
C:Date: 21-Dec-1996 #sequence_revision 06-Jun-1997 #text_change 16-Dec-1998
C:Accession: G01025
R:Navarro, E.
submitted to the EMBL Data Library, April 1996
A:Reference number: H00564
A:Accession: G01025
A>Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-745 <NAV>
A:Cross-references: EMBL:X97630; NID:g1310674
C:Superfamily: unassigned Ser/Thr or Tyr-specific protein kinases; protein kinase homolog
F:18-27/Domain: protein kinase homolog <kin>

Query Match 21.6%; Score 687.5; DB 2; Length 745;
Best Local Similarity 28.1%; Pred. No. 2e-21;
Matches 202; Conservative 108; Mismatches 225; Indels 183; Gaps 21;

Qy 2 KVEREIALKLEIHPHVLKLDVYENKYLVLVLEHVSQGLFDYLVKKGRLTTPKEARKF 61
Db 63 KLFREVRIMKVLNHPNIVKLFVEVETETKTLVLMVEYASGGEVDFYLVAGHGRMKEARAK 122
Qy 62 FQRIISALDFCHSHSICHRLDKPENLLDKNINRIADFGMASLQVGDLSLETSCGSPHY 121
Db 123 FQIIVSAVQYCHQKFIHVHRLDKAENLLDADNMNIKIDFGFSNEFTFGNKLDTFCGSPY 182
Qy 122 ACPRVIRGEKYDGRKADVWSGVLFPALLVGCALPFDNDNLRQLLEKVKRGVFMHPPIPP 181
Db 183 AAPELFGKLYDGEVDVWSGVLFPALLVGCALPFDNDNLRQLLEKVKRGVFMHPPIPP 242
Qy 182 DCQSLRGMIEVDAARLLTLEHIQKHIWYIGGKNEPEPEQPIPRKVQIRSLPSLEDDID 241
Db 243 DCENLLKFLILNPSKRGTLQIMKDRMVMVGHEDDE-----LKPVEPLPDYK---DPR 294
Qy 242 VLDSMHSIGCFRDRNKLQDLSEENQEKMIYFLLLDRKERYPSQEDDILPRNEID-- 299
Db 295 RTELVMWNGVYTRBE---IQSLVGQRYNEYMATYLLGLYKSSLEGGDTITLKPRPSADLT 351
Qy 300 -----PPRKVDSD-----PMLNRHGK-----RRPE-----R 320
Db 352 NSSAQFPESHKQSVSNANPKQRFSDQAGPAITPSNSYKKTQSNNAENKRPEDRESGR 411
Qy 321 KSM-----EVLVYVDG---GSPVPARRAIEMA---QHQR 350
Db 412 KASSTAKVASPLPGLERKKTITPTPSTNSVLSTSTNRNSPLLRASLQASIQNGKDS 471
Qy 351 RSTSG---ASSGLSTSPSSPRVTPH-----PSPRGSPLTPKGTVPHT 391
Db 472 LTMFGSRASASAGVSAARPRQHOKSMGSGVHPNKAAGLPPTESNCEVREPTAPQV 531
Qy 392 PKESPA-----GTNPNT-----PPSSVGVG 413
Db 532 PVASPSAHNITSSSGCAPDRTNPRGVSSRSTFHAGQLRQVRDQQLPYGVTPASPSGHSQ 591
```

[illegible]

Db 726 TEFO 729

RESULT 8

T52633

serine/threonine-specific protein kinase (EC 2.7.1.1-) AKIN11 [validated] - Arabidopsis
N:Alternate names: SNF1 protein kinase homolog AKIN11
C:Species: Arabidopsis thaliana (mouse-ear cress)
C>Date: 24-Oct-2000 #sequence_revision 24-Oct-2000 #text_change 03-Nov-2000
C:Accession: T52633

R:Bhalerao, R.P.; Salchert, K.; Bako, L.; Okresz, L.; Szabados, L.; Muranaka, T.; Machida
Proc. Natl. Acad. Sci. U.S.A. 96, 5322-7, 1999
A>Title: Regulatory interaction of PRU1 WD protein with Arabidopsis SNF1-like protein ki
A:Reference number: 225116; MUID:99238528; PMID:10220464
A:Accession: T52633
A>Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: mRNA

A:Residues: 1-512 <BA>

A:Cross-references: EMBL:X99279; PIDN:CAA67671.1

A:Experimental source: cultivar Columbia

C:Genetics:

A:Gene: AKIN11

C:Function:

A:Description: EC 2.7.1.-; serine/threonine-specific protein kinase AKIN11 [validated, M
complements SNF1 mutations in yeast

C:Superfamily: AMP-activated protein kinase; protein kinase homolog

C:Keywords: ATP; phosphotransferase; serine/threonine-specific protein kinase

Query Match 21.4%; Score 691.5; DB 2; Length 512;

Best Local Similarity 39.5%; Pred. No. 2.4e-21;

Matches 140; Conservative 73; Mismatches 106; Indels 35; Gaps 6;

Qy 2 KVEREIAIKLIEHPHVLKLDVYENKYLVLVLEHVSGLGELFDYLVKKGRLTPKEARKF 61

Db 64 KVRERIKLRLFMHPIRQYEVLETSIDIVVMEYVKSGLFDYLVKKGRLQDEARNF 123

Qy 62 FROIISALDFCHSHSICHRDLKPNLLDEKNIRIADFGMASLQVGDLSLLETSCGSPHY 121

Db 124 FQIISGVYCHRNVMVHRDLKPNLLDSKCNKIADFGLSNVRDGHFLKTS CGSPNY 183

Qy 122 ACPEVIRGEKYDGRKADVNSGVILFALLVGLALPDDNLRQLLEKVKRGVFMHPHIPP 181

Db 184 AAPEVISGLYAGPEVDVMSGVILYALLCGLTPFDENIPNLFKKGITLPSHLS 243

Qy 182 DCQSLLRGMEVDAAARLTLSHIQKHIWYIGGKNEPEPEQIPRKVOIRSLPSLED--I 238

Db 244 EARDLIPRLMIVDPVKRITIEIRQHRNF-----QTHLPRLVAVSPDTPVEQAKI 294

Qy 239 DPVDLSMHSGLCFDRNKLQDLSEENQEKMIYFLLDRKERYPSQEDDLPPENEI 298

Db 295 NEEIVQEVNMGF--DRNQVLESRLNRQTNDATVTVYLLDNRRPVSGLY-----SEF 347

Qy 299 DPKRKVDSPMLNRHGRKPERKSNVLSVTDGSPVPARAEIAEQHQRSS 352

Db 348 QETDSSGNP-----RTPEA-----GASPVGHWIPAHVDHYGLGARS 385

RESULT 9

S59941

serine/threonine-specific protein kinase (EC 2.7.1.1-) BKIN2 - barley (fragment)

C:Species: Hordeum vulgare (barley)

C>Date: 15-Feb-1996 #sequence_revision 01-Mar-1996 #text_change 26-Feb-1999

C:Accession: S59941

R:Hannappel, U.; Vicente-Carabajosa, J.; Barker, J.H.A.; Shewry, P.R.; Halford, N.G.

Plant Mol. Biol. 27, 1235-1240, 1995

A>Title: Differential expression of two barley SNF1-related protein kinase genes.

A:Reference number: S59941; MUID:95284374; PMID:7766906

A:Accession: S59941

A>Status: preliminary

A:Molecule type: mRNA

A:Residues: 1-473 <BA>

A:Cross-references: EMBL:X82548

C:Genetics:

A:Gene: BKIN2

C:Function:

A:Description: catalyzes the formation of peptidyl-serine-phosphate or peptidyl-threonine
C:Superfamily: AMP-activated protein kinase; protein kinase homolog
C:Keywords: ATP; magnesium; phosphotransferase; serine/threonine-specific protein kinase
F;1-244/Domain: protein kinase homolog (fragment) <KIN>
F;21,40,115,117/Active site: Lys, Glu, Asp, Lys #status predicted
F;120,124/Binding site: magnesium (Asn, Asp) #status predicted

Query Match 21.1%; Score 672.5; DB 1; Length 473;

Best Local Similarity 42.7%; Pred. No. 5.3e-21;

Matches 128; Conservative 69; Mismatches 82; Indels 21; Gaps 4;

Qy 2 KVEREIAIKLIEHPHVLKLDVYENKYLVLVLEHVSGLGELFDYLVKKGRLTPKEARKF 61

Db 36 KVRERIKLRLFMHPIRQYEVLETSIDIVVMEYVKSGLFDYLVKKGRLQDEARNF 95

Qy 62 FROIISALDFCHSHSICHRDLKPNLLDEKNIRIADFGMASLQVGDLSLLETSCGSPHY 121

Db 96 FQIISGVYCHRNVMVHRDLKPNLLDSKCNKIADFGLSNVRDGHFLKTS CGSPNY 155

Qy 122 ACPEVIRGEKYDGRKADVNSGVILFALLVGLALPDDNLRQLLEKVKRGVFMHPHIPP 181

Db 156 AAPEVISGLYAGPEVDVMSGVILYALLCGLTPFDENIPNLFKKGITLPSHLS 215

Qy 182 DCQSLLRGMEVDAAARLTLSHIQKHIWYIGGKNEPEPEQIPRKVOIRSLPSLEDI 238

Db 216 LARDLIPRLMIVDPVKRITIEIRHSWF-----KARLPRLVAVPPDTPAQVKKL 266

Qy 239 DPVDLSMHSGLCFDRNKLQDLSEENQEKMIYFLLDRKERYPS-----QEDD 291

Db 267 DDETLDNVKMGF--DKNQVLESRLNRQTNDATVTVYLLDNRRPVSGLYAEYQESMD 324

RESULT 10

S59359

GIN4 protein - yeast (Saccharomyces cerevisiae)

N:Alternate names: protein YDR507c

C:Species: Saccharomyces cerevisiae

C>Date: 30-Nov-1995 #sequence_revision 01-Mar-1996 #text_change 24-Sep-1999

C:Accession: S59359; S59565

R:Longtine, M.S.; Pringle, J.R.

submitted to the EMBL Data Library, August 1995

A:Reference number: S59359

A:Accession: S59359

A:Molecule type: DNA

A:Residues: 1-1142 <ION>

A:Cross-references: EMBL:U33140; NID:G992650; PIDN:AAA75513.1; PID:G992651

R:Dietrich, F.S.

submitted to the EMBL Data Library, August 1995

A:Description: The sequence of S. cerevisiae cosmids 8166, 9787, 9717, and lambda 3073.

A:Reference number: S59553

A:Accession: S59565

A:Molecule type: DNA

A:Residues: 1-1142 <DIE>

A:Cross-references: EMBL:U33057; NID:G927764; PIDN:AAB64949.1; PID:G927777; MIPS:YDR507c

C:Genetics:

A:Gene: SGD:GIN4

A:Cross-references: SGD:S0002915; MIPS:YDR507c

A:Map position: 4R

C:Superfamily: unassigned Ser/Thr or Tyr-specific protein kinases; protein kinase homolog

C:Keywords: ATP

F;17-289/Domain: protein kinase homolog <KIN>

F;25-33/Region: protein kinase ATP-binding motif

Query Match 21.0%; Score 668.5; DB 2; Length 1142;

Best Local Similarity 34.1%; Pred. No. 1.8e-20;

Matches 183; Conservative 73; Mismatches 162; Indels 119; Gaps 16;

Qy 3 VEREIAIKLIEHPHVLKLDVYENKYLVLVLEHVSGLGELFDYLVKKGRLTPKEARKF 62

Db 78 IEREIIMKLLNHPNVLRYDVWETNTDLYLVLEYAEKGLFNLLVERGRLPSHEAIRFF 137

QY 63 RQIIISALDFCHSHSICHRLDKPENLLDDEKNNIRIADFGMASLQVDSGLLETSCGSPHYA 122
Db 138 RQIIIGVSYCHALGIVHRDLKPENLLDHDKNYIKIADFGMALETETGKJLETSCGSPHYA 197
QY 123 CPEVIRGEKDGKADVWSCGVILFALLVGLALPPD--DNLRLLEKVKRGVHMP--HF 178
Db 198 APEIVSGIPYCGFASDVWSCGVILFALLVGLALPPDDEEDGIRTLILLVKVKGSEFMPDDE 257
QY 179 IPPDCQSLRGMIIEVDAAARLTLEHIQKHIWYIGGKNEPEPEPIPKV-----QIRS 231
Db 258 ISREAQDLIRKILTVDERIKTRDIKH-----PQKQYPSIRDSKSIKRG 303
QY 232 LP-----SLEDIDPDVLSMHSGLCFDRNKLQDLSEENQKMIYFLLDLR 280
Db 304 LPREDTYLPLSENSSIDATILQNLVILWHGRDPEGIEKJLREPAGANAETLYALLY-- 361
QY 281 KERYPSQDEDLPRNEIDPRKRVDSFPLNRHGRKRPKRSMEVLSVTDGSGVPARRA 340
Db 362 --RPFKCDTQKELIKQQV-----KXQQSISVSVS-----PSKV 394
QY 341 IEMAQHGQRSISGASGLSTSPSGPRVTPHPSRGSPLETPKGTPTVHTPKESPAGTP 400
Db 395 STTPQRRNRRESLSVTS-----RKQPSIFNKFT-----ASSASS 431
QY 401 NPTPPSPSVGVPWRLNRIKNSIFLGSPPFRHKLQVTPPEMSNLTPESSPELAKKS 460
Db 432 NLTPFGSK-----RL-----SKNFSSK-----KLSTIVNQSSPTPASRN 467
QY 461 WFGNFISLEKEEQ---IFVVIKDKPLSLKADIVHAFSLTSPSLSHSVISOTSFAEY 514
Db 468 KRASVINVEKNQKRASIFSTTKNKRSSRS---IKRMSLIPSMKRESVT-TKLMSTY 520

RESULT 11
A56009
serine/threonine-specific protein kinase (EC 2.7.1.-) NPXS - common tobacco
C;Species: Nicotiana tabacum (common tobacco)
C;Date: 28-Apr-1995 #sequence_revision 28-Apr-1995 #text_change 16-Jun-2000
C;Accession: A56009
R;Muranaka, T.; Banno, H.; Machida, Y.
Mol. Cell. Biol. 14, 2958-2965, 1994
A;Title: Characterization of tobacco protein kinase NPXS, a homolog of Saccharomyces cerevisiae.
base of Saccharomyces cerevisiae.
A;Reference number: A56009; MUID:94217693; PMID:8164654
A;Accession: A56009
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-511 <MUR>
A;Cross-references: GB:D26602; NID:G496384; PIDN:BAA05649.1; PID:G496385
C;Function:
A;Description: catalyzes the formation of peptidyl-serine-phosphate or peptidyl-threonine
C;Superfamily: AMP-activated protein kinase; protein kinase homology
C;Keywords: ATP; magnesium; phosphotransferase; serine/threonine-specific protein kinase
F;17-271/Domain: protein kinase homology <KIN>
F;25-33/Region: protein kinase ATP-binding motif
F;48,67,142,144/Active site: Lys, Glu, Asp, Lys #status predicted
F;147,151/Binding site: magnesium (Asn, Asp) #status predicted

Query Match 21.0%; Score 668; DB 1; Length 511;
Best Local Similarity 44.4%; Pred. No. 8.7e-21;
Matches 128; Conservative 61; Mismatches 85; Indels 14; Gaps 3;

QY 2 KVEREIALKLEIHPHVLKLDVYENKYLVLVLEHVSGLGELFDYLVKKGLRTPKEARKF 61
Db 63 KVRREIKILRFMHPHIIIRYEVVETPSDIYVWVEYKSGELFDYVVEKGLQDEARKF 122
QY 62 FROIIISALDFCHSHSICHRLDKPENLLDDEKNNIRIADFGMASLQVDSGLLETSCGSPHY 121
Db 123 FQOIIISGVEYCHRNWVHRDLKPENLLDSDKNWVKVIADFGLSNIRMDGHFLTKSCGSFNY 182
QY 122 ACPEVIRGEKDGKADVWSCGVILFALLVGLALPPDDNLRLLEKVKRGVHMPHPTPP 181
Db 183 AAPEVISGKLYAGPEVDVWSCGVILYALLCGTLPPDDENIPNLFKKIKGMISLPSHLSA 242

QY 182 DCQSLRGMIIEVDAAARLTLEHIQKHIWYIGGKNEPEPEPIPKVQI---RSLPSLEDI 238
Db 243 GARDLIPMLIIVDPKRMWIPERHHPWF-----QAHLPRYLAVPPPTMQQAKKI 293
QY 239 DPVLDLSMHSGLCFDRNKLQDLSEENQKMIYFLLDLRKYRPS 286
Db 294 DEDILQEVVKGRF--DRNSLVASLCNRVQNEGTAVAYLILENQFRASS 339

RESULT 12
T41587
probable carbon catabolite derepressing protein kinase - fission yeast (Schizosaccharomyces pombe)
C;Species: Schizosaccharomyces pombe
C;Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 31-Jan-2000
C;Accession: T41587
R;Murphy, L.; Harris, D.; Lyne, M.; Rajandream, M.A.; Barrell, B.G.
submitted to the EMBL Data Library, September 1998
A;Reference number: Z21968
A;Accession: T41587
A;Status: preliminary; translated from GB/EMBL/DBJ
A;Molecule type: DNA
A;Residues: 1-576 <MUR>
A;Cross-references: EMBL:AL031543; PIDN:CAA20833.1; GSPDB:GN00068; SPDB:SPCC74.03c
A;Experimental source: strain 972h; cosmid c74
C;Genetics:
A;Gene: SPDB:SPCC74.03c
A;Map position: 3
A;Intons: 50/1; 127/3; 245/1
C;Superfamily: AMP-activated protein kinase; protein kinase homology

Query Match 20.8%; Score 662.5; DB 2; Length 576;
Best Local Similarity 37.3%; Pred. No. 1.6e-20;
Matches 146; Conservative 76; Mismatches 114; Indels 55; Gaps 10;

QY 1 MKVEREIALKLEIHPHVLKLDVYENKYLVLVLEHVSGLGELFDYLVKKGLRTPKEARK 60
Db 77 MRVEREISYLLKLRHPIIKLYDVIPTTPTDIVMVEY-AGGELFDYVVEKGRWTEDEGR 135
QY 61 FPROIISALDFCHSHSICHRLDKPENLLDDEKNNIRIADFGMASLQVDSGLLETSCGSPH 120
Db 136 FQOIIICAEYCHRHKIVIRDLKPENLLDLDNVLKVIADFGLSNIRMDGHFLTKSCGSPN 195
QY 121 YACPEVIRGEKDGKADVWSCGVILFALLVGLALPPDDNLRLLEKVKRGVHMPHPTIP 180
Db 196 YAAPEVINGKLYAGPEVDVWSCGVILYVLMVGLRPFDDDFIPNLFKKVNSCVYVMPDFLS 255
QY 181 PDCQSLRGMIIEVDAAARLTLEHIQKHIWYIGGKNEPEPEPIPKVQIERSLPSLEIDP 240
Db 256 PGAQSLIRRMIVADPMQRTIQEIRDPWF--NVNLPDYLRP-----MEEVQG 301
QY 241 DVLDSDM-----HSLGCFDRNKLQDLSEENQKMIYFLLDLRKYRPSQEDDLPP 294
Db 302 SYADSRIVSKLGEAMGFSED--YIVELASDENNEVKEAYNLL--HENQVIOEKSHL-- 354
QY 295 RNEIDPPRKRVDS-----PMLNRHGRKRPKRSMEVLSVTDG-----SPVPARRAI 341
Db 355 -----SKSKRVDSFLSVSPAPSEYTSLEQKSKQLIDPTLEGPRWTWSDPTTAKQTI 409
QY 342 E-----MAHQGQRSISGASGLSTS 363
Db 410 DSNICVLVPTAEKNKLEMRMTLADAASAVDTS 440

RESULT 13
I48609
probable serine/threonine-specific protein kinase (EC 2.7.1.-) kem - mouse
C;Species: Mus musculus (house mouse)
C;Date: 02-Jul-1996 #sequence_revision 02-Jul-1996 #text_change 04-Mar-2000
C;Accession: I48609; S31333
R;Ingalls, J.D.; Lee, M.; Hill, R.E.
Mamm. Genome 4, 401-403, 1993
A;Title: Emk, a protein kinase with homologs in yeast maps to mouse chromosome 19.

A;Reference number: I48609; MUID:93364122; PMID:8358177

A;Accession: I48609

A;Status: preliminary; translated from GB/EMBL/DBJ

A;Molecule type: mRNA

A;Residues: 1-774 <ING1>

A;Cross-references: EMBL:X70764; NID:957919; PIDN:CAA5040.1; PID:957920

Ringlis, J.D.; Lee, M.; Hill, R.E.

submitted to the EMBL Data Library, January 1993

A;Description: A novel protein kinase with homologues in yeast maps to mouse chromosome

A;Reference number: S31333

A;Accession: S31333

A;Molecule type: mRNA

A;Residues: 1-698, 'K', 700-702, 'GRGLPTAAKASCTGCTCYCAGNAHQATRT', 731-774 <ING2>

A;Cross-references: EMBL:X70764

C;Genetics:

A;Gene: emk

C;Superfamily: unassigned Ser/Thr or Tyr-specific protein kinases; protein kinase homolog

C;Keywords: ATP; phosphotransferase; serine/threonine-specific protein kinase

F:51-304/Domain: protein kinase homology <XIN>

F:59-67/Region: protein kinase ATP-binding motif

Query Match 20.7%; Score 661.5; DB 2; Length 774;
Best Local Similarity 27.2%; Pred. No. 2.4e-20;
Matches 195; Conservative 111; Mismatches 227; Indels 183; Gaps 19;

QY 2 KVEREIALKLEHHPVHLKLDVYENKYLVLVLEHVGSGELFDYLVKKGRLTPKEARKF 61

Db 96 KLFREVIMKVLNHNIVKLFVETETKTLVLMVYASGGEVDFVLVAHGRMKEAKAK 155

QY 62 FQIISALDFCHSHSICHRLDKPENLLDEKNIRIADFGMASLQVDSLETSCGSPHY 121

Db 156 FQIIVLHVQYCHQKFIIVHRLDKAENLLDADNIRIADFGFSGNEFTFGNKLDTCGSPPY 215

QY 122 ACPEVIRGEKYDGRKADVWSCGVILFALLVGLPDDNLRQLLEKVKRGVFMHPHTPP 181

Db 216 AAPELFQKKGKIDGPEVWLSGLVILVLSGLPFGQNLRELRLVLRGKYRIPFYNST 275

QY 182 DCQSLRGMIEVDAAARLTLEHQHIVYIGKNEPEPEQIPRKVQIRSLPSLEDIDPD 241

Db 276 DCENLKKFLILNPSKRGTEQIMKDRWMNVGHEDELKPYVE-----PLLTGPRD 327

QY 242 VLDSMHSGLCFRDRNKLQDLSSSEENQRMKIFLLDRKRYPSQEDDLPPEIDPD 299

Db 328 RVDGVNGL-----HTEEIQDSLVGRYNEVMATYLLGLYSSEPEGDITLKPRPSADLT 382

QY 300 -----PPRKVDVS-----PMLNRHGK-----RPERK----- 321

Db 383 NSSAPSPSHKVQSVSANPKORSSDQAVPAIPTSNSYSKKTQSNNAENKRPSEETGRKA 442

QY 322 -----SMEVLSVTDG---GSPVPARRAIEMA---QHGQGRS 352

Db 443 SSTAKVPASPLPGLDRKKTTPAPSTNSVLSTNSRNSPLLDRLASLGQSIQNGKSLT 502

QY 353 ISG---ASSGLSTSPSSPRVTPH-----PSRGSPLPTPKGTPTVHTPK 393

Db 503 MPGRASASTASAAVAAPRQHQKMSBASVHPNKAASGLPTESNCEVPEPSTAPQRPV 562

QY 394 ESP-----AGTPNPT-----SSSRHASMSMP-----QKPAIRGRTTSSMP 415

Db 563 ASPSMNLISSSGADPTNFPGRVSSRTTFIAGQLQRVDRQONLPYGVTPASPGHSGQR 622

QY 416 EARLNSIKNSFLGSPRRHRLQVPTPEMS---NLTPSSPELAKKSWFGNFISL 468

Db 623 RGASGIFSKF--TSKFVRNLNPESEKDRVETLRLPHVVSGGTDKKEEFREAKPRSLR 680

QY 459 KSW-FGNFISLEKEHQIFVWIKDPLSISKADIVHAFSLGISLSHSVTSQTSFRAEYKAT 517

Db 681 FTWSMKTTSSENEPMEIRKVLDAQCSELHERYMLL--CVHGTGPHENF----- 731

QY 518 GGPVAFQKPVKFOVDITTEGEAQKNGIYSVFTLLSGPSRRFKRVVETIQAL 573

Db 732 -----VOWEMEVC-----KULPRSLNGVRFRKIRISGTSMAFKNIASIANEL 772

RESULT 14

T29858

hypothetical protein T01C8.1 - *Caenorhabditis elegans*

C;Species: *Caenorhabditis elegans*

C;Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 18-Feb-2000

C;Accession: T29858

R;Wohldmann, P.; Hawkins, J.

submitted to the EMBL Data Library, May 1996

A;Description: The sequence of *C. elegans* cosmid T01C8.

A;Reference number: Z20699

A;Accession: T29858

A;Status: preliminary; translated from GB/EMBL/DBJ

A;Molecule type: DNA

A;Residues: 1-562 <MOH>

A;Cross-references: EMBL:U58726; PIDN:AAS00579.1; GSPDB:GN00028; CESP:T01C8.1

C;Experimental source: strain Bristol N2; clone T01C8

C;Genetics:

A;Gene: CESP:T01C8.1

A;Map position: X

A;Introns: 41/1; 197/2; 237/1; 272/2; 307/3; 375/1; 479/1; 529/3

C;Superfamily: AMP-activated protein kinase; protein kinase homology

Query Match 20.7%; Score 660.5; DB 2; Length 562;
Best Local Similarity 30.6%; Pred. No. 1.9e-20;
Matches 164; Conservative 93; Mismatches 158; Indels 121; Gaps 15;

QY 2 KVEREIALKLEHHPVHLKLDVYENKYLVLVLEHVGSGELFDYLVKKGRLTPKEARKF 61

Db 69 KIRREIQLSLFRPHIIRLYQVISTPDSIFIMIEHVSGGELFDYLVKGRKLTAEARRF 128

QY 62 FQIISALDFCHSHSICHRLDKPENLLDEKNIRIADFGMASLQVDSLETSCGSPHY 121

Db 129 FQIISGVDYCHRMVVRDLKPENLLDEQNNVKIADFGLSNIMTGDGLFRTSCGSPNY 188

QY 122 ACPEVIRGEKYDGRKADVWSCGVILFALLVGLPDDNLRQLLEKVKRGVFMHPHTPP 181

Db 189 AAPEVISGKLYAGPEVDVWSCGVILYALLCGTLFPDDHVPFLFRKIKSGVFPPTDFLER 248

QY 182 DCQSLRGMIEVDAAARLTLEHQHIVYIGKNEPEPEQIPRKVQIRSLPSLEDIDPD 241

Db 249 PIVNLLEHMLCVDPMKRAITKDVIAHEWF-----QKDLPNYL-----PFPINESEAS 295

QY 242 VLDSMHSGLCFRDRNKLQD-----LLSEENQKMI-YELLDRK----- 281

Db 296 IVD-IEAVREVTREYHVAEEVTSALLGDDPHHLSIAYNLIVDNKRIADETAKLSTEEP 354

QY 282 -----BRYPSQEDDLPPEIDPPRRKRVDSPLNRH-----GKRPE 319

Db 355 YQVTPNKGPGVHRHPERIAASV--SSKITPTLDNTEASGANRNKRAKWHLGIRSQSRPE 412

QY 320 RKSNEVLVS-----TDGSPVPARRAIEMAHQGRS-----RSI 353

Db 413 DIMPEVFRAMKQDMKWLNPYHVIIVRRKPDAPADPPKMSLQLYQVDORSLYLLDFKSL 472

QY 354 SGASSGLSTSPSSPRVTPHPSRGSPLPTPKGTPTVHTPKESPAGTNPPTPPSPSVGGV 413

Db 473 ADRESGSASA-----SSSRHASMSMP-----QKPAIRGRTTSSMP----- 508

QY 414 PWRLNSIKNSFLGSPRRHRLQVPTPEMS---NLTPSSPELAKKSWFGNFISL 468

Db 509 ----QAMSMEASI-----EKMEVHDFSDMSCDVTPPPSPGAKLSQTMQFPEI 552

RESULT 15

S74283

probable protein kinase YCLO24w (BC 2.7.1.1) - yeast (*Saccharomyces cerevisiae*)

C;Species: *Saccharomyces cerevisiae*

C;Date: 19-Mar-1997 #sequence_revision 25-Apr-1997 #text_change 19-Apr-2002

C;Accession: S74283; S19351; S19741

R;Voet, M.; Voelckert, G.

submitted to the Protein Sequence Database, September 1996

A;Reference number: S74277

A:Accession: S74283
A:Molecule type: DNA
A:Residues: 1-915 <VOE>
A:Cross-references: ENBL:X59720, NID:gl907116; PIDN:CAA42361.1; PID:e309035; PID:gl90712
R:Duisterhoft, A.; Erdmann, D.; Hegemann, J.; Philippsen, P.; Schweitzer, B.; Spiegelberg
submitted to the Protein Sequence Database, March 1992
A:Reference number: S19351
A:Accession: S19351
A:Molecule type: DNA
A:Residues: 1-569 <DUS>
A:Cross-references: ENBL:X59720; MIPS:YCL024w
R:Olliver, S.G.; Anwar, R.; Brown, A.; Gent, M.E.; Indge, K.J.; James, C.M.; Stateva, L.L.
submitted to the Protein Sequence Database, March 1992
A:Reference number: S19337
A:Accession: S19741
A:Molecule type: DNA
A:Residues: 567-803, 'SRRWHIT', 807-808, 'RH', 811, 'G' <OLI>
A:Cross-references: ENBL:X59720; MIPS:YCL024w
A:Note: this sequence has been revised in reference S74283
C:Genetics:
A:Gene: SGD:KCC4
A:Cross-references: SGD:S0000529
A:Map position: 3L
A:Note: YCL024w
C:Superfamily: unassigned Ser/Thr or Tyr-specific protein kinases; protein kinase homolog
C:Keywords: ATP; phosphotransferase; serine/threonine-specific protein kinase
F:19-285/Domain: protein kinase homology <KIN>
F:27-35/Region: protein kinase ATP-binding motif

Query Match 20.5%; Score 655; DB 2; Length 915;
Best Local Similarity 34.2%; Pred. No. 5.2e-20;
Matches 177; Conservative 73; Mismatches 177; Indels 90; Gaps 13;

Qy 3 VERETAILKLEHPHVLLKHDVYENKYLVLVHVSAGELPFDYLVKKGRTPKEARKFF 62
Db 74 IEREIVIMKLLSHPNVLSYDWTNNLLYLLEYAEKGEFLNLLVDHGPDPERAINCF 133
Qy 63 RQIIISALDFCHSHSICHKDLKPELLEKNNIRIADFGMASLQVGDLSILETSCGSPHYA 122
Db 134 RQIIIGISVCHALGIVHRDLKPELLEKNNIRIADFGMAALQTDADLLETSCGSPHYA 193
Qy 123 CPREVIRGEKYDGRKADWSCGVILFALLVGLPFDDDD--NLRLLEKVRGVFHMPPH--F 178
Db 194 APEIVSGLPFYEGFASDVWSCGVILFALLTGRLPFDEENGVRDLKLVKQGFEMPNDTE 253
Qy 179 IPPDQCSLLRGMIEVDAAARLLTLEHIQHIWIYIGGKNEPEQPIPRKVO--IRSLPSLED 237
Db 254 ISRDAQDLIGKILVDPQRIRKIRDLSH-----PLKKYQITIDSKSIND 299
Qy 238 -----IDPDVLDMSHSLGCFRDRNKLLOLLSEENQEKMIYFL 276
Db 300 LPRENTLYPLADSNHTSASIDDSILQNLVVLWHRHADDIVSKLKENGNTNKEKILYAL 359
Qy 277 LLDRKERYPSQDEDLPPRNEIDPPRKVRVDSPLNLRHKGKRRPKRSMVLSVTDGSPVP 336
Db 360 LY--RFKLDVSRGSKNKNKNIKTKKNKRSTLSSSSLLLNRSIQ-----STP 408
Qy 337 ARRAATMAHQGORSISGASSGLSTSPSSPVPHPSP--RGSPLTPKGPVHTPKES 395
Db 409 RRRTSK-----RHSREFS--SSKRSGFLSSNPTDSSPILRSSKRITHINVASANTQAT 462
Qy 396 PAGTRNPTPPSPSPVGGVFWRLNLSIKNSFLGSPFRHRKLQVPTPEMSNLTPESSPE 455
Db 463 PSGVFNP-----HKRNSKRSKSLSYMPNTKRSS 492
Qy 456 LAKSNFGNFISLEKEEQIFVVIKD--KPLSSIKADIV 491
Db 493 LTKSKS--LSNFTNLIDDDWEYIEKDAKRTSSNFATLI 528

Search completed: November 26, 2003, 12:02:39

Job time : 17.6771 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 26, 2003, 11:50:16 ; Search time 10.0063 Seconds
(without alignments)
2857.432 Million cell updates/sec

Title: US-10-054-579-4

Perfect score: 3188

Sequence: 1 MKVREITAIKLIEHPVLK.....TNCMMMTGRLSKGIIPKS 608

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 127863 seqs, 47026705 residues

Total number of hits satisfying chosen parameters: 127863

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_41:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	722	22.6	1518	1 KKK1 YEAST	P34244 saccharomyc
2	698	21.9	512	1 K110 ARATH	Q38997 arabidopsis
3	685.5	21.5	775	1 CDR2 SCHPO	P87050 schizosacch
4	675.5	21.2	752	1 MRK4 HUMAN	Q96134 homo sapien
5	668.5	21.0	1142	1 G1N4 YEAST	Q12263 saccharomyc
6	662.5	20.8	576	1 SNF1 SCHPO	O74536 schizosacch
7	661.5	20.7	774	1 MRK2 MOUSE	Q05512 mus musculu
8	655	20.5	915	1 KCC4 YEAST	P25389 saccharomyc
9	645	20.2	786	1 SN1L HUMAN	P57059 homo sapien
10	638	20.0	776	1 SN1L RAT	Q97155 rattus norv
11	637	20.0	776	1 MRK3 HUMAN	P27448 homo sapien
12	636	19.9	552	1 AAK2 RAT	Q09137 rattus norv
13	632	19.8	552	1 AAK2 HUMAN	P54646 homo sapien
14	631.5	19.8	548	1 AAK1 RAT	P54645 rattus norv
15	631	19.8	779	1 SN1L MOUSE	Q60670 mus musculu
16	627.5	19.7	550	1 AAK1 HUMAN	Q13131 homo sapien
17	620.5	19.5	619	1 SNF1 CANTR	O94168 candida tro
18	616	19.3	633	1 SNF1 YEAST	P06782 saccharomyc
19	611.5	19.2	502	1 RK11 SECCCE	Q02723 secale cere
20	609.5	19.1	611	1 SNF1 CANAL	Q03372 candida gla
21	596.5	18.7	620	1 SNF1 CANAL	P52497 candida alb
22	585.5	18.4	891	1 KIN1 SCHPO	P22987 schizosacch
23	556	17.4	622	1 YNA3 CAEEL	P45894 caenorhabdi
24	546.5	17.1	1064	1 KIN1 YEAST	P13185 saccharomyc
25	546	17.1	661	1 Y537 HUMAN	O60285 homo sapien
26	533	16.7	593	1 CDR1 SCHPO	P07334 schizosacch
27	500	15.7	1147	1 KIN2 YEAST	P13186 saccharomyc
28	493	15.5	800	1 KIN4 YEAST	Q01919 saccharomyc
29	461.5	14.5	353	1 ASK2 ARATH	P43292 arabidopsis
30	458.5	14.4	363	1 ASK1 ARATH	P43291 arabidopsis
31	447	14.0	714	1 HUNK HUMAN	P57058 homo sapien
32	444.5	13.9	714	1 HUNK MOUSE	O88866 mus musculu
33	433.5	13.6	664	1 KCCB HUMAN	Q13554 homo sapien

RESULT 1

```
KKK1 YEAST
ID _KKK1 YEAST STANDARD; PRT; 1518 AA.
AC P34244;
DT 01-FEB-1994 (Rel. 28, Created)
DT 01-FEB-1994 (Rel. 28, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE Probable serine/threonine-protein kinase YKL101W (EC 2.7.1.1.-).
GN YKL101W OR YKL453.
OS Saccharomyces cerevisiae (Baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
OX NCBI_TaxID=4932;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=S288C;
RX MEDLINE=94078677; PubMed=8256524;
RA Pallier C., Valens M., Fuzos V., Fukuhara H., Cheret G., Sor F.,
RA Bolotin-Fukuhara M.;
RT "DNA sequence analysis of a 17 kb fragment of yeast chromosome XI
RT physically localizes the MRB1 gene and reveals eight new open reading
RT frames, including a homologue of the KIN1/KIN2 and SNF1 protein
RT kinases."
RL Yeast 9:1149-1155(1993).
CC -!- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
CC NIM1 SUBFAMILY.
-----
This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
-----
EMBL; X71133; CAA50456.1; -.
EMBL; Z28101; CAA81941.1; -.
PIR; S37928; S37928.
HSSP; Q63450; 1A06.
SGD; S0001584; HSL1.
GO; GO:0005935; C:bud neck; IDA.
GO; GO:0005940; C:septin ring; IDA.
GO; GO:0000086; P:G2/M transition of mitotic cell cycle; IGI.
GO; GO:0006468; P:protein amino acid phosphorylation; IDA.
GO; GO:0000074; P:regulation of cell cycle; IMP.
GO; GO:0000135; P:septin checkpoint; IGI.
InterPro; IPR000719; Prot_kinase.
InterPro; IPR002290; Ser_Thr_kinase.
Pfam; PF00069; kinase; 1.
PRINTS; PR00109; TYRKINASE.
ProDom; PD000001; Prot_kinase; 1.
SMART; SM00220; S_TKC_1.
PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
-----
O14012 homo sapien
O91y88 mus musculu
Q63450 rattus norv
Q35555 homo sapien
Q02066 triticum ae
O15831 homo sapien
O75385 homo sapien
O13557 homo sapien
P15791 rattus norv
P28178 dictyosteli
P11730 rattus norv
Q923t9 mus musculu
```

ALIGNMENTS


```
DR PROSITE; PS00108; PROTEIN KINASE ST; 1.
KW Transferrase; Serine/threonine-protein kinase; ATP-binding;
FT DOMAIN 10 262 PROTEIN KINASE (BY SIMILARITY).
FT NP_BIND 16 24 ATP (BY SIMILARITY).
FT ACT_SITE 133 133 BY SIMILARITY.
FT BINDING 39 39 ATP (BY SIMILARITY).
SQ SEQUENCE 775 AA; 85971 MW; 2EE2EBCA8C84FAC CRC64;

Query Match 21.5%; Score 685.5; DB 1; Length 775;
Best Local Similarity 29.3%; Pred. No. 3e-28;
Matches 212; Conservative 99; Mismatches 210; Indels 203; Gaps 26;

Qy 2 KVEREIALKLIHPHVLKDVYENKYLVLVLEHVSQGLFVYLVKKGRLTPKEARKP 61
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 62 FQIISALDFCHSHSICHRDLKPNLLDEKNNIRIADFGMASIQVGDLSLETSCGSPHY 121
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 114 LWQILCGLEYCHKLHICHRDLKPNLYLDAGHSIKIGFGMASIQQPKLLQTSQSPHY 173
Qy 122 ACPEVIREKYGDKADYVSCGVLFPALLVGLPDDNLRQLLEKVKRGVFNHPHIPP 181
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 174 ASPEIIMGRSYDGCASDIWSCGIIFPALLTGKLPDDNIRSLKLVQCGQFEMPSPNIS 233
Qy 182 DCQSLLRGMIEVDARRLLTLEHIOKHWIYGGKNEP-----EPEQIPRKVQIRSLPSL 235
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 234 QAQHLLYMLVDSDSTRIMQIREH-----PFLSCFVHNISIP-----IISAP-I 279
Qy 236 EDIDPDVLDSMHL--GCFDRNRKLLQDLLSEENQEKMIYFLLLDRKERYPSQDEDLPP 294
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 280 QPIDPLIVQHLSLVPRCDDPMPLEKASQSPLEKTLTYLL--SRHLHP-----PS 330
Qy 295 RNEIDPPKRVDSPM--LNRHCKRPERKSEVLVSVTGGSPVPARRAIEAQQGQRS 352
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 331 SAAVDNRNAVVDLLGTAAASGQQQWDEEIEQAINIPT-LAPYPISYAAESVPRPATAS 389
Qy 353 -----ISGA--SSGLSTPLSSPRVTPHSPRGSPLP-TPKG-TP-VHTPKESPACTP 400
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 390 PFLTPTVTSGTFNYSFNATNQS---ILQRATTSSAVPQLPKSVTPGLAYPHDSMLSS 446
Qy 401 NPTPPSS-----PSV-----GGV 413
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 447 NYRPPSALSPRNFNVSINDPEVQLSRRATSLDMSNDFRMNENDPSVGNLAASNPPTGMG 506
Qy 414 PWRARLNSIKNSFLG-----SPRFR-----434
Qy 507 PPRKVTGRMSEHTGNRVVSPRGSFAFNPRTVRFVNGNEQFSNNIDNNNYQPYANATMN 566
Qy 435 --RKLQVTPTE-----EMSNLTPE-----SPELAKKSW 461
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 567 NSRRLRTSGERSMRADLSQSPASYDSLNVPKHRRRQSLFSPSTTKKLSGSPQPKESP 626
Qy 462 -----FGNFTSLKEKQIFVVIKDKPLSSIK-ADIVHAFSLTSPSLSHSVIS 506
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 627 LRRFLSSEPPCKCVVASLVSLEHLEILVLRWQLLGIGADIYDVS-ASISARIKR 685
Qy 507 QTSFRAEYKATGPAVFOKPKVQFOVDITVTEGEAQKNGIYSVTFTLLSGPSRRFKVW 566
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 686 QNSLNL-----KPVFRISVLAEFPGS-----QAVFVLESSTTFDHLA 725
Qy 567 ETIQ 570
Db 726 TEFG 729

RESULT 4
MRK4 HUMAN
ID MRK4 HUMAN STANDARD; PRT; 752 AA.
AC Q96L34; Q96JG7; Q9YB8;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
```

```
DR EMBL; AK075272; BAC11510.1; --
DR EMBL; AB058763; BAB47489.1; ALT_INIT.
DR HSSP; O63450; 1A06.
DR Genew; HGNC:13538; MARK4.
DR MIM; 606495; -
DR InterPro; IPR001772; Kinase_Cterm.
DR InterPro; IPR000719; Prot_Kinase.
DR InterPro; IPR002290; Ser_Thr_pkinase.
DR InterPro; IPR001245; Tyr_pkinase.
DR InterPro; IPR000449; UBA_domain.
DR Pfam; PF02149; KAI; 1.
DR Pfam; PF00069; pkinase; 1.
DR Pfam; PF00627; UBA; 1.
DR PRINTS; P000109; TYRKINASE.
DR ProDom; P000001; Prot_Kinase; 1.
DR PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
DR PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
DR PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
DR PROSITE; PS50030; UBA; 1.
DR Transferase; Serine/threonine-protein kinase; ATP-binding;
KW Alternative splicing.
FT DOMAIN 59 310 PROTEIN_KINASE.
FT DOMAIN 324 368 UBA.
FT DOMAIN 703 752 KAI.
FT NP_BIND 65 73 ATP (BY SIMILARITY).
FT BINDING 88 88 ATP (BY SIMILARITY).
FT ACT_SITE 181 181 BY SIMILARITY.
FT VARSPIC 628 752
PEPEALMALQNTAAARCRKQPPQFLACLHGAGGPEPL
SHEFEVCLPRGLRGVLFRRVAGTALAFRTLVTRISNDL
EL -> TLDPKRONRCVSGASLPQSGKIRSQTNLRSG
DLRSQVATYLGIRKRPKPPGCSDSGV (in isoform
2).
FTID=VSP 004946.
FT CONFLICT 70 70 F -> S (IN REF. 2).
FT SEQUENCE 752 AA; 82519 MW; 4B430FDD2B150E7A CRC64;

Query Match          21.2%; Score 675.5; DB 1; Length 752;
Best Local Similarity 29.3%; Pred. No. 9.5e-28;
Matches 205; Conservative 94; Mismatches 223; Indels 177; Gaps 22;

QY 2 KVERIAIKLIEHPVHLKHDVYENKYLVLVHVGSGELFDVLVKKGLRTPKEAKRF 61
Db 102 KLFREVRIMKGLNHNIVKLFVETIKTLVLMEYASAGEVDFVLVHGMRKEARAK 161
QY 62 FROILSALDFCHSHICHRLDKPENLLDKKNIRIADFGMASLVQGSLLTSCGSPHY 121
Db 162 FQIIVSAVHYCHQKNIVHRDLKAEILLDAENIKIADFGFSNEFTGLSKLDTFCGSPY 221
QY 122 ACPEVIRGEYDKRKYADVMSCCVILFALLVGLPFDNDNLRLEKVKRGVFMHPHTPP 181
Db 222 AAPELFQGGKYDGPEDVINSGLVILYLVSGSLPFDGHNKLRERVLKRGYVPFYMYST 281
QY 182 DCQSLLRGMIEVDARRLTLEHIQKHW----YIGGKNP--EPQPIPRKVQIRSLPSL 235
Db 282 DCESILRFLVLPNPAKRTLEQIMKDKMINIGYEGLKPYTEP----- 326
QY 236 EDI-PPDVLDSMHSIGCFDRNKLQDLISEENQEKMIYFLLDKERYPSQEDEDLP 294
Db 327 EDFGDKRIEVMVGYYTKEIK--ESLTSQKYNVTATY--LLGRK-----TEEGD--- 376
QY 295 RNEIDPPKRVDSPLNLRHG----KRRPERKSMELVSVT-----DGGSPVPARRA 340
Db 377 RGAPGLALARVAPDVTNGTSSSGKTSKSGQRSSSTYHQRHSDFCGSPAPLHPK 436
QY 341 IMAHQGQ-----RSRISGASGLSTPLSSRPVTPHPSPSPPLTPKGTVPVH 390
Db 437 RSPTSTGEALKEERLPGRKASCSTAGSGRLPLPSPSPVMSAHPNKAETIPERRKDS 496
QY 391 TP-----KESPACTNPPTPPSPS-----VGG 412
Db 497 TPNNLPPSMTRNTYVCTERPGAERPSLLPNKNGNSSGTPR--VPPASPSHSLAPPSGE 555
```

```
QY 413 VPMRRLNLSIKNSFLG-----SPRFHRRKLQVPT--PEEMSNL 448
Db 556 RSLRARGSTIRSFHGGQVDRDRAGGGGGVQNGPPASPFLAHEAAPLPAGRPRTN 615
QY 449 -----TPSSPELAKKSWFGNFISLEKEBQIFV 476
Db 616 FTKLTSKLTRRVADEPRIGGPEVTSCHLEPMDQETATAPRLLRFPWSVKLTSSRPEALMA 675
QY 477 VIKDKPLSS-IKADIVHAFSLPSLSHVSISQTSFRAEYKATGPGVAFQKPVK-FOVDIT 534
Db 676 ALRQATAAACRCRQPPFL-----LACHGGAGGP-----EPLSHFEVEVC 717
QY 535 YTEGGEAKENGIVSVTFTLLSGPSRRFRKRVETIQAL 573
Db 718 QL-----PRPLRGVLFRRVAGTALAFRTLVTRISNDL 750

RESULT 5
GIN4_YEAST
ID_GIN4_YEAST STANDARD; PRT; 1142 AA.
AC Q12263;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE Serine/threonine-protein kinase GIN4 (EC 2.7.1.1.-).
GN GIN4 OR YDR507C OR D9719.13.
OS Saccharomyces cerevisiae (Baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
OX NCBI_TaxID=4932;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=99030835; Pubmed=9813093;
RA Longtine M.S., Fares H., Pringle J.R.;
RT "Role of the yeast Gln4 protein kinase in septin assembly and the
RL relationship between septin assembly and septin function.";
RL J. Cell Biol. 143:719-736(1998).
RN [2]
RP SEQUENCE FROM N.A.
RA Dietrich F.S., Mulligan J., Allen E., Araujo R., Aviles E., Berno A.,
RA Carpenter J., Chen E., Cherry J.M., Chung E., Duncan M.,
RA Hunnicke-Smith S., Hyman R., Komp C., Lashkari D., Lew H., Lin D.,
RA Mosedale D., Nakahara K., Namath A., Oefner P., Oh C., Petel F.X.,
RA Roberts D., Schramm S., Schroeder M., Shogren T., Shroff N.,
RA Winant A., Yelton M., Botstein D., Davis R.W.;
RL Submitted (AUG-1995) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: MAY PLAY A ROLE IN SEPTIN ASSEMBLY.
CC -!- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
CC NIM1 SUBFAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; U33140; AAA75513.1; -.
CC EMBL; U33057; AAB64949.1; -.
CC PIR; S59359; S59359.
CC HSSP; Q63450; 1A06.
CC SGD; S0002915; GIN4.
CC GO; GO:0005935; C:bud neck; IDA.
CC GO; GO:0007117; P:bud growth; IGI.
CC GO; GO:0000094; P:septin assembly and septum formation; IGI.
CC GO; GO:0001135; P:septin checkpoint; IGI.
CC InterPro; IPR000719; Prot_kinase.
CC InterPro; IPR002290; Ser_Thr_kinase.
CC InterPro; IPR001245; Tyr_kinase.
CC Pfam; PF00069; pkinase; 1.
CC PRINTS; P000109; TYRKINASE.
CC ProDom; P000001; Prot_kinase; 1.
```

DR SMART; SMO0220; S TKC; 1.
 DR PROSITE; PS00107; PROTEIN KINASE ATP; 1.
 DR PROSITE; PS00108; PROTEIN KINASE ST; 1.
 DR PROSITE; PS00011; PROTEIN KINASE DOM; 1.
 KW Transferase; Serine/threonine-protein kinase; ATP-binding.
 FT DOMAIN 19 289
 FT NP BIND 25 33 ATP (BY SIMILARITY).
 FT BINDING 48 48 ATP (BY SIMILARITY).
 FT ACT SITE 156 156 BY SIMILARITY.
 SQ SEQUENCE 1142 AA; 129857 MW; EC16P4B849DD811 CRC64;
 Query Match 21.0%; Score 668.5; DB 1; Length 1142;
 Best Local Similarity 34.1%; Pred. No. 3.5e-27;
 Matches 183; Conservative 73; Mismatches 162; Indels 119; Gaps 16;
 QY 3 VEREIAIKLIEHPVHLKHDVYENKYLVLVLEHVSGLFDYLVKKGRLTPKEARKFP 62
 DB 78 IEREIIMKLNHPNVLRLDVTETNDLVLEAYAEKGLFNLLVERGLPHEAIRFP 137
 QY 63 RQIISALDFCHSHSICHRLDKPENLLDEKNNIRIADFGMASLQVGSLLLETSCGSPHYA 122
 DB 138 RQIIGVSYCHALGIVHRDLKPENLLDHKNYKIADFGMAALETEGKULETSCGSPHYA 197
 QY 123 CPVIRGEKDGKADVWSCGVILFALLVGLALPPD--DDNLROLLEKVRGVFHP--HF 178
 DB 198 APEIVSGIPYQGFASDVWSCGVILFALLTGRLPDEEDGNIRTLKLVQKGEPEMPDDE 257
 QY 179 IPPDCQLLAGMIEVDAAERLTLEHIOKHIWYIGKNEPEPEQIPKV-----QIRS 231
 DB 258 ISREAQLIRKILVDFERIKTRDLKH-----PLIQKPSIRDSKISRG 303
 QY 232 LP-----SLEIDIPDLDMSHSLGCFDRNKLQDLASEENQEKMIYFLLDR 280
 DB 304 LPREDTYLPLSESSSIDATILQNVLVHGRDPEGIKLEPGANAETKLYALLY-- 361
 QY 281 KERYPSQEDDLPPRNEIDPPKRVDSFMLNHRHGRKPRKSMVLSVTGSGSPVARRA 340
 DB 362 --RPKCDTQKELIKQQV-----KRGQSSSVSVS-----PSKKV 394
 QY 341 IEMAHQGRSISGASGLSTSLSPRVTPHPSRGSPLTPKGTVPVHTPKESPAGTP 400
 DB 395 STTPQRNRRESLSVTS-----RKPFISFNKFT-----ASSASS 431
 QY 401 NPTPPSPFSGVGPWRARLNSIKNSFLGSPRFRHKLQVTPPEMSNLTPESSPELAKS 460
 DB 432 NLTPGSSK-----RL-----SKNFSKK-----KLSTVQSSFTPASRN 467
 QY 461 WFGNFISLEKEQ---IFVVIKQPLSSIKADIVHAFSLPSLSHVSISQTSFRAEY 514
 DB 468 KRSVINVERKQKRSISFSTTKQKRSRS---IKRMSLIPSMKRESVT-TKLMSTY 520
 RESULT 6
 SNF1_SCHPO STANDARD; PRT; 576 AA.
 AC OY4536;
 DT 15-JUL-1999 (Rel. 38, Created)
 DT 15-JUL-1999 (Rel. 38, Last sequence update)
 DT 15-SEP-2003 (Rel. 42, Last annotation update)
 DE SNF1-like protein kinase (EC 2.7.1.1-).
 GN SPOC74.03C.
 OS Schizosaccharomyces pombe (Fission yeast).
 OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
 OC Schizosaccharomycetales; Schizosaccharomycetaceae;
 OC Schizosaccharomycetes.
 OX NCHI_TaxID=4896;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=972;
 RC MEDLINE=21848401; PubMed=11859360;
 RA Wood V., Gwilliam R., Rajandream M.A., Lyne M., Lyne R., Stewart A.,
 RA Sources J., Feat N., Hayles J., Baker S., Basham D., Bowman S.,
 RA Brooks K., Brown D., Brown S., Chillingworth T., Churcher C.M.,

RA Collins M., Connor R., Cronin A., Davis P., Feltwell T., Fraser A.,
 RA Gentles S., Goble A., Hamlin N., Harris D., Hidalgo J., Hodgson G.,
 RA Holroyd S., Hornsby T., Howarth S., Huckle E.J., Hunt S., Jagels K.,
 RA James K., Jones L., Jones M., Leather S., McDonald S., McLean J.,
 RA Mooney P., Moule S., Mungall K., Murphy L., Niblett D., Odell C.,
 RA Oliver K., O'Neil S., Pearson D., Quail M.A., Rabinowitsch E.,
 RA Rutherford K., Rutter S., Saunders D., Seeger K., Sharp S.,
 RA Skelton J., Simmonds M., Squares R., Squares S., Stevens K.,
 RA Taylor K., Taylor R.G., Tivey A., Walsh S.V., Warren T., Whitehead S.,
 RA Woodward J., Volckaert G., Aert R., Robben J., Grymonprez B.,
 RA Welljens I., Vanstreels E., Rieger M., Schaefer M., Mueller-Auer S.,
 RA Gabel C., Fuchs M., Fritzc C., Holzer E., Moestl D., Hilbert H.,
 RA Borzym K., Langer I., Beck A., Leirach H., Reinhardt R., Pohl T.M.,
 RA Eger P., Zimmermann W., Wedler H., Wambutt R., Purnelle B.,
 RA Goffeau A., Cadieu E., Dreano S., Gloux S., Lelaure V., Mottier S.,
 RA Galibert F., Aves S.J., Xiang Z., Hunt C., Moore K., Hurst S.M.,
 RA Lucas M., Rochet M., Gallardin C., Tallada V.A., Garzon A., Thode G.,
 RA Daga R.R., Cruzado L., Jimenez J., Sanchez M., del Rey F., Benito J.,
 RA Dominguez A., Revuelta J.L., Moreno S., Armstrong J., Forburg S.L.,
 RA Cerrutti L., Lowe T., McCombie W.R., Paulsen I., Potashkin J.,
 RA Shpakovski G.V., Ussery D., Barrell B.G., Nurse P.;
 RL "The genome sequence of Schizosaccharomyces pombe";
 Nature 415:871-880(2002).
 CC -1- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
 CC -1- SUBFAMILY.
 CC -1- SIMILARITY: Contains 1 UBA domain.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See http://www.isb-sib.ch/announce/
 or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; AL031543; CAA20833.1; --
 DR F1R; T41587; T41587.
 DR HSSP; O63450; 1A06.
 DR GeneDB SPombe; SPOC74.03C; --
 DR InterPro; IPR000719; Prot_kinase.
 DR InterPro; IPR002230; Ser_thr_kinase.
 DR InterPro; IPR000449; UBA_domain.
 DR Pfam; PF00069; pkinase; 1.
 DR ProDom; PD000001; Prot_kinase; 1.
 DR SMART; SMO0220; S TKC; 1.
 DR PROSITE; PS00107; PROTEIN KINASE ATP; 1.
 DR PROSITE; PS00011; PROTEIN KINASE DOM; 1.
 DR PROSITE; PS00108; PROTEIN KINASE ST; 1.
 DR PROSITE; PS00030; UBA; 1.
 KW Transferase; Serine/threonine-protein kinase; ATP-binding.
 FT DOMAIN 34 285
 FT NP BIND 304 345 UBA.
 FT BINDING 63 63 ATP (BY SIMILARITY).
 FT ACT SITE 156 156 BY SIMILARITY.
 SQ SEQUENCE 576 AA; 65996 MW; E5857E8F171E7B50 CRC64;
 Query Match 20.8%; Score 662.5; DB 1; Length 576;
 Best Local Similarity 37.3%; Pred. No. 3.2e-27;
 Matches 146; Conservative 76; Mismatches 114; Indels 55; Gaps 10;
 QY 1 MKVEREIAIKLIEHPVHLKHDVYENKYLVLVLEHVSGLFDYLVKKGRLTPKEARK 60
 DB 77 MRVEREISVLKLRHPIIKLVYITPTDVMVIEY-AGGELFDYIVKRGKRWTEDEGR 135
 QY 61 FFRQIISALDFCHSHSICHRLDKPENLLDEKNNIRIADFGMASLQVGSLLLETSCGSPH 120
 DB 136 FFRQIISALDFCHSHSICHRLDKPENLLDEKNNIRIADFGMASLQVGSLLLETSCGSPH 195
 QY 121 YACPEVIRGEKDGKADVWSCGVILFALLVGLALPPD--DDNLROLLEKVRGVFHPMFHIP 180
 DB 196 YAAPEVINGKLYAGPEVDVWSCGVILVLMVGLPFDDEDFIPLNFKVNSCVVYMPDFLS 255


```

QY 181 PDCQLRGMIEVDAAARLTLEHIOKHIWYIGGKNEPEQPIPRKVOIRSLPSLEDIDP 240
Db 256 PQAQLIRMIADVADPMQRTIOEIRDPWF--NVNLPDYLRP-----MEEVQG 301
QY 241 DVLDSM-----HSLGCFRDRNKLQDLISEBENOEKMIYFLLDRKERYPSQDEDLPP 294
Db 302 SYADSRIVSKLGEANGFSED--YIVAEALSDENNEVEKAYNLL---HENQVIOEKSHL-- 354
QY 295 RNEIDPPKRVDS-----PMLNRHGKRRPERKSMIEVLSVTDG-----SPVPAARAI 341
Db 355 -----SKSRVDSFLSVSPAPASEYTSSELOKKSQGLIDPTLEGPRWTVPDPPTVAKQTI 409
QY 342 E-----MAHQGQRSISGASGLSTS 363
Db 410 DSNICVLPTAEKNKLEMKRTLAASAVIDTS 440

RESULT 7
ID_MRK2_MOUSE STANDARD; PRT; 774 AA.
AC Q05512;
DT 01-FEB-1995 (Rel. 31, Created)
DT 01-FEB-1995 (Rel. 31, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE MAP/microtubule affinity-regulating kinase 2 (EC 2.7.1.27)
DE (Serine/threonine-protein kinase Emk).
GN MARK2 OR EMK.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RX TISSUE=Embryo;
RC MEDLINE=93364122; PubMed=8358177;
RA Inglis J.D., Lee M., Hill R.E.;
RT "Emk, a protein kinase with homologs in yeast maps to mouse
RT chromosome 19."
RL Mamm. Genome 4:401-403 (1993).
CC -!- CATALYTIC ACTIVITY: ATP + a protein = ADP + a phosphoprotein.
CC -!- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
CC MARK SUBFAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; X70764; CAA50040.1; --
CC PIR; I48609; I48609.
CC HSP; O63450; 1A06.
CC MGD; MGI:99638; Mark2.
CC InterPro; IPR001772; Kinase_Cterm.
CC InterPro; IPR000719; Prot_Kinase.
CC InterPro; IPR002290; Ser_Thr_kinase.
CC InterPro; IPR001245; Tyr_kinase.
CC Pfam; PF02149; KAI; 1.
CC Pfam; PF00069; pkinase; 1.
CC PRINTS; PR00109; TYRKINASE.
CC ProDom; PD000001; Prot_kinase; 1.
CC SMART; SM00220; S_TK; 1.
CC PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
CC PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
CC PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
CC Transferrase; Serine/threonine-protein kinase; ATP-binding.
KW DOMAIN 53 304 PROTEIN_KINASE.
FT NP_BIND 59 67 ATP (BY SIMILARITY).
FT BINDING 82 82 ATP (BY SIMILARITY).
FT ACT_SITE 175 175 BY SIMILARITY.
SQ SEQUENCE 774 AA; 85874 MW; 02BF8D7BF443483A CRC64;

```

```

Query Match 20.7%; Score 661.5; DB 1; Length 774;
Best Local Similarity 27.2%; Pred. No. 5.1e-27;
Matches 195; Conservative 111; Mismatches 227; Indels 183; Gaps 19;

QY 2 KVEREIALKLIHPIVHLKLDHYVYENKYLVLVLEHVHVGGLLEFDYLVKKGRLTKPEAKRF 61
Db 96 KLFREVIMKVLNPNIVKULFEVETETKLYLMVEYASGGVEFDFVLVAHGMKKEARAK 155
QY 62 FROIISALDFCHSHSICHROLKPNLLILDEKNNIRIADFGMASLQVGSLLLETCGSPHY 121
Db 156 FROIIVLHVQCHQKFIHVHRLKAENLLLDADMNLIKADFGSNEFTFGNKLDTFCGSEPPY 215
QY 122 ACPEVIRGEKYGDKKADVMWCGVILFALLVGALEPDDNLRQLLEKVKRGVFNHPIPP 181
Db 216 AAPLFOGKKIDGPEVDVWMLGVLTYLVSGSLPFDGQNLKELRERVLRGKYRPFYFMST 275
QY 182 DCQSLRGMIEVDAARLTLEHIOKHIWYIGGKNEPEQPIPRKVOIRSLPSLEDIDP 241
Db 276 DCENLLKKFLILNPSKRGTLLEQIMKDRMNVGHDELKPYVE-----PLTTGPRD 327
QY 242 VLSMHSLGCFRDRNKLQDLISEBENOEKMIYFLLDRKERYPSQDEDLPPNEID-- 299
Db 328 RVDGVNGL-----HTEIQDSLVCQRYNEVMATYLLGLYKSSSEPGDTITLKPRPSADLT 382
QY 300 -----PPRKRVDS-----PMLNRHGK-----RRPERK----- 321
Db 383 NSSAPSPSHKQVRSVANPKQRRSSDQAVPAIPTSNSYSKKTQSNNAENKPEETGRKA 442
QY 322 -----SMEVLVTDG---GSPVPARRAEMA---OHGQRSRS 352
Db 443 SSTAKVPASPLFGLDRKKTTPAPSTNSVLSTNRSNPLDRLASLGQASIQNGKOSLT 502
QY 353 ISG--ASSGLSTSLSPSRVTPH-----PSPRGSLPTPKGTPTVHTPK 393
Db 503 MPGSEASTASASAASAAARPRHQKMSASVHPNKAQSLPTESNCEVPRSTAPQRPV 562
QY 394 ESP-----AGTPNPT-----PPSSPSVGGVPM 415
Db 563 ASPSAHNISSSGAPDRTNFPRGVSSRSTFHAGQLQRVDRQDNLPGVTTPASPSGHSQGR 622
QY 416 RARLNSIKNSPLGSPFRHRRKLOVP--TPEEMSLTP-----ESSPELAK 458
Db 623 RGASGSIFSKP--TSKVRRLNLEPESKDRVETURPHVVGSGGTDKDKKEPRKAPSLR 680
QY 459 KSW--FGNFISLEKEEQFVVIKDKPLSSIKADIVHAFILSPSLSHSVISQTSFRAEYKAT 517
Db 681 FTWSMKTSSNEPNEEMREIRKVLDSANSCQSELHERYMLL--CVHGTPTGHENF----- 731
QY 518 GGPAVQKPKVQVDITYTTEGGEAKENGIVSVTFTLLSGSPRRPKRVVETIQAL 573
Db 732 -----VQWEMEVC-----KLPRLSLNGVRFKRISGTSMAFKNIASKIANEL 772

RESULT 8
ID_KCC4_YEAST STANDARD; PRT; 915 AA.
AC P25389; P87005;
DT 01-MAY-1992 (Rel. 22, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE Probable serine/threonine-protein kinase YCL24W (EC 2.7.1.-).
GN YCL024W OR YCL24W.
OS Saccharomyces cerevisiae (Baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
OX NCBI_TaxID=4932;
RN [1]
RP SEQUENCE OF 1-569 FROM N.A.
RA Duesterhoeft A., Erdmann D., Hegemann J., Philippsen P.,
RA Schweitzer B., Spiegelberg R.;
RL Submitted (MAR-1992) to the EMBL/GenBank/DBJ databases.
RN [2]

```

RP SEQUENCE OF 567-915 FROM N.A.
 RA Oliver S.G., Anwar R., Brown A., Gent M.E., Indge K.J., James C.M.,
 RA Staveva L.I.,
 RL Submitted (MAR-1992) to the EMBL/GenBank/DBJ databases.
 RN [3]
 RP REVISIONS.
 RA Gromadka R.,
 RL Submitted (JAN-1996) to the EMBL/GenBank/DBJ databases.
 CC -!- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
 CC NIML SUBFAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

CC EMBL; X59720; CAA42361.1; -.
 DR PIR; S74283; S74283.
 DR HSSP; O63450; 1A06.
 DR SGD; S0000529; YCL024W.
 DR GO; GO:0005935; Cbud neck; IDA.
 DR GO; GO:0003700; F:transcription factor activity; IDA.
 DR GO; GO:0007117; P:bud growth; IGI.
 DR GO; GO:0000094; P:septin assembly and septum formation; IGI.
 DR GO; GO:0000135; P:septin checkpoint; IGI.
 DR InterPro; IPR000719; Prot_kinase.
 DR InterPro; IPR002230; Ser_thr_kinase.
 DR InterPro; IPR001245; Tyr_pkinase.
 DR Pfam; PF00069; pkinase; 1.
 DR PRINTS; PR00109; TYRKINASE.
 DR ProDom; PD000001; Prot_kinase; 1.
 DR SMART; SM00220; S_TKC_1.
 DR PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
 DR PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
 DR PROSITE; PS00111; PROTEIN_KINASE_DOM; 1.
 KW Hypothetical protein; Transferase; Serine/threonine-protein kinase;
 KW ATP-binding.
 FT DOMAIN 21 285 PROTEIN KINASE.
 FT NP_BIND 27 35 ATP (BY SIMILARITY).
 FT BINDING 50 50 ATP (BY SIMILARITY).
 FT ACT_SITE 152 152 BY SIMILARITY.
 SQ SEQUENCE 915 AA; 102688 MW; BFB01C8CA43AC181 CRC64;

Query Match 20.5%; Score 655; DB 1; Length 915;
 Best Local Similarity 34.2%; Pred. NO. 1.3e-26;
 Matches 177; Conservative 73; Mismatches 177; Indels 90; Gaps 13;
 3 VERETAILKLEHPHVLKLDHYENKKYLYLVLEHVGGEFDYLVYKGRLPFKEARKFF 62
 74 IREIVIMKLLHPNVLSDYVWETNNLYLILEVAEKGEFLNLLVDHGPFPREAINCF 133
 63 ROIISALDFCHSHSICRDLKPEKLLDKNRIRADFOMASLQVDSLETSCSPHYA 122
 134 ROIIGISYCHALGIVHRLKPEKLLDFFYAKIADFOMALQTDADLETSCSPHYA 193
 123 CPEVIRGEYKGRKADWCGVILFALLVGPFDGDD--NLRLLEKVKRGVFMH--F 178
 194 APEIVSGLYEYFASVWCGVILFALLVGRLPFDDENGVNDLLKVGKGFEMPNDE 253
 179 IPPDCQSLRGMIEVDAARRLTLEHQIHWIGGKNEPEPEQIPRKVQ--IRSLPSLED 237
 254 ISRDAQDLGKILVDPQRKIRKIDLSH-----PLKKYQTIKDSKIKD 299
 238 -----IDPDVLDMSHSLGCFRDNKLLQDLLEENQEKMYPL 276
 300 LPRENTYLPADSNHNTGASIDDSILQNLVLMHGRHADDIVSKLKENGNTKELIAL 359
 277 LLDRKERYPSQEDDELPNNEIDPPRKRYVDSPLNLRHGRKRPKRKMEVLSDVGGSPVP 336
 360 LY--RFKLDVSGSNKKNRNKIKTKTKNKRSTLSSSSILLNNRSIQ-----STP 408

QY 337 ARRAIEAQAQGORSRSGISGASGLSTSPVTPHPSP-RGSPLPKGTVPVHTPKES 395
 DB 409 RRTSK-----RHSREFS--SRKRSFLSSNTDSSPIPLRSKRITHINVASANTQAT 462
 QY 396 PAGTPNTPPSPSPSVGVWRAKLNSIKNSIFLGSFRRKQLQVTPPEMNNITPSSSPE 455
 DB 463 PSQVNP-----HKNSKRSKRSKRLSYMPNTPKRSS 492
 QY 456 LAKSMFGNFISLEKEBQIFWIKD-KPLSSIKADIV 491
 DB 493 LTSKS-LSNFTNLDDDDWEYIEKDAKRTSSNFATLI 528

RESULT 9

ID SNIL_HUMAN STANDARD; PRT; 786 AA.
 AC P57059;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 15-SEP-2003 (Rel. 42, Last annotation update)
 DE Probable serine/threonine protein kinase SNF1LK (EC 2.7.1.-).
 GN SNF1LK.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=20289795; PubMed=10830953;
 RA Hattori M., Fujiyama A., Taylor T.D., Watanabe H., Yada T.,
 RA Park H.-S., Toyoda A., Ishii K., Totoki Y., Choi D.-K., Groner Y.,
 RA Soeda E., Ohki M., Takagi T., Sakaki Y., Taudien S., Blechschmidt K.,
 RA Polley A., Menzel U., Delabar J., Kumpf K., Lehmann R., Patterson D.,
 RA Reichwald K., Rump A., Schillhabel M., Schudy A., Zimmermann W.,
 RA Rosenthal A., Kudoh J., Shibuya K., Kawasaki K., Asakawa S.,
 RA Shantani A., Sasaki T., Nagamine K., Mitsuyama S., Antonarakis S.E.,
 RA Minoshima S., Shimizu N., Nordsiek G., Hornischer K., Brandt P.,
 RA Schafhe M., Schoen O., Desario A., Reichelt J., Kauer G., Bloeker H.,
 RA Ramser J., Beck A., Klages S., Hennig S., Rieselmann L., Dagand E.,
 RA Wehrmeyer S., Borzym K., Gardiner K., Mizetic D., Francis F.,
 RA Lehrach H., Reinhardt R., Yaspo M.-L.; -."
 RL "The DNA sequence of human chromosome 21."
 RL Nature 405:311-319(2000).
 CC -!- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
 CC SNF1 SUBFAMILY.

CC -!- SIMILARITY: Contains 1 UBA domain.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

CC EMBL; AP001751; BAA95536.1; -.
 DR HSSP; P24941; 1A01.
 DR Genew; HGNC:11142; SNF1LK.
 DR MIM; 605705; -.
 DR InterPro; IPR000719; Prot_kinase.
 DR InterPro; IPR002290; Ser_thr_kinase.
 DR InterPro; IPR001245; Tyr_pkinase.
 DR InterPro; IPR000449; UBA domain.
 DR Pfam; PF00069; pkinase; 1.
 DR PRINTS; PR00109; TYRKINASE.
 DR ProDom; PD000001; Prot_kinase; 1.
 DR SMART; SM00220; S_TKC; 1.
 DR SMART; SM00165; UBA; 1.
 DR PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
 DR PROSITE; PS00111; PROTEIN_KINASE_DOM; 1.
 DR PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
 DR PROSITE; PS00030; UBA; 1.

```
KW Transferase; Serine/threonine-protein kinase; ATP-binding.
FT DOMAIN 27 281
FT NP BIND 306 346
FT BINDING 33 41
FT ACT SITE 56 56
FT ACT SITE 149 149
SQ SEQUENCE 786 AA; 85252 MW; FB44EE7CAF87CB1A CRC64;

Query Match
Best Local Similarity 20.2%; Score 645; DB 1; Length 786;
Matches 203; Conservative 84; Mismatches 238; Indels 134; Gaps 24;

QY 2 KYREIAIKLIEHPHVLKLDHYENKYLVLVLEHVSQGLFDYLVKKGRLTPKEARKF 61
DB 70 KYREVQLMKLLNHPNIIKLYQVMEKMDLYIVTEFAKNGEMFDYLTNGHLSNEARKK 129
QY 62 FQIISALDFCHSHSTCHRDLPENLLDEKNIRIA---DFGMASLQVGDLSLSTSGS 118
DB 130 FQWILSAVEYCHDHHIVHRDLKTENLLDGNMDIKLAGTDFGNGFYKSGEPLSTWCS 189
QY 119 PHYACPEVIRGEKYGKADVMSQGVLFALLVGLPFDGDDNLRQLLEKVRGVFHPHF 178
DB 190 PYYAAPEVEGEKYEYEGPQDLSGLVGLVGLVGLPFDGPNLPTLRQVLEGRFRIPFF 249
QY 179 IPPDCOSLRGMIEVDAAARLTLEHIQKHIWIGGKNEPEPQIPRKVQIRSLPSLEDI 238
DB 250 MSQDCESLIRMLVDPARITIAQIRQHRWRAEPCLPACPAFSAHSYTS---NLGDY 307
QY 239 DPVDLDSMHSGLCFDRNKLLOLLSEENQEMKIYFLLLDKERY-----P 285
DB 308 DQALGIMQTLGV--DRQTVESLQNSYNHFAIYLLERLKEYRNAQCARPGARQP 365
QY 286 SQEEDLP---PRNEI---DP-----PRKRVDSPM----- 309
DB 366 RPRSSDLGLEVPQEGSLSTDPFPALLCPQPTLVQSVLQAEWDCQLQSLQWPLFPVD 425
QY 310 LNRHGKRRERKSMVSLVTDGSPVPARR--AIENAHQGRS-----RSTSGASSGLS 361
DB 426 ASCSGVFRP--RPVSSSLDDTAISEARQGPGLSEEDQTSPLSPSTGRRHTLAEVSTR 483
QY 362 TSPLSRPVTPHPSRPGSLPTPKGTP-----VHTPKESPAGTPNTPSPSVGV-PW 415
DB 484 LSPLTAPCKFVSPSTASP---AEGTSSDSCUTFSASKSPAGL-SGTPTQGLGACSPV 539
QY 416 RARLNSIKNSFLGSPFRHRKQLQVTPPEKMSNLTPSSPELAKKWFNGFISLEKEEQIF 475
DB 540 R-----LSPFLGS-----QSATPVLAQGGGLGAVLLPVSFQSG 574
QY 476 VVIKDPKLS-----IKADIVHAF-----SIPSLSHVISQTSFRAEYKATGCPAVF 523
DB 575 RRASDTSLTQGLKAFRQQLKTRTKGFLGNLKNKIGLARQVCQAPASRA---SRGSLSPF 631
QY 524 QKPVKQVDITTEGGEAOKENGIYSVTFTLLSGPSRRFRKRVVETIQALLSTHDPAA 582
DB 632 HAPAQ-----SPGLHGGAAGSREG-----WSLL-----BEVLE--QORLLQLOHPAA 672

RESULT 10
SNIL RAT
ID SNIL RAT STANDARD; PRT; 776 AA.
AC Q9RIU5; Q9R081;
DT 16-OCT-2001 (Rel. 40, Created)
DT 15-SEP-2003 (Rel. 42, Last sequence update)
DE Probable serine/threonine protein kinase SNFLK (EC 2.7.1.-) (Salt-
DE inducible protein kinase) (Protein kinase KID2).
GN SNFLK OR SIK OR KID2.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
```

Db 363 RQPOLRNDLSLEVPQBIPLCDPFRPSLPCFPQALAQSLQAEICDLHSLSLQPLFFP 422
Qy 319 -----ERKSEVLVSVTGGSPVPARRAIEAQAQGRSISGASSGLST----- 362
Db 423 LDTWCVGVRHRSISPSLLDTAISEARQPSLEEEQVEPFGTGRHTLAEVSTH 482
Qy 363 -SPSSRVTPHPSPRGSPPLTPKGT-----VHTPKESPAG-----TNPTPPSPSPV 410
Db 483 FSLNPPCCIIVSSAAVSP-----SEGTSDSCCLPFSASEGPAAGLGGGLATPGLGTSSP-- 537
Qy 411 GGVPWRRLNLSIKNSFLGSPFRHKKLVQVPEMSN-----LTPESSE----- 455
Db 538 -----VRLAS-----PFLGS-----QSATPVIQSQAAGLGTALVPPVSFQGRASDTS 580
Qy 456 -----LAKSWFGNFISLEKEQIFWIKDKPLSLIKADIVHAFILSIPSLSHSV 504
Db 581 LTQGLKAFRQRLKRNARKYKGLNKK-----INGLARQV 614
Qy 505 ISQTSFRAYKATGPAVFPQK 526
Db 615 -QSSSIRG---SRGGMSTFHTP 632

RESULT 11
MRK3_HUMAN
ID MRK3_HUMAN STANDARD; PRT; 776 AA.
AC P27448; O60219; Q8TB41; Q8WX83; Q96RG1; Q9UMY9; Q9UN34;
DT 01-AUG-1992 (Rel. 23, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE MAP/microtubule affinity-regulating kinase 3 (EC 2.7.1.27) (Cdc25C-associated protein kinase 1) (CTAK1) (C-TAK1) (Serine/threonine protein kinase p78) (Ser/Thr protein kinase PAR-1) (Protein kinase STK10).
DE MARK3 OR CTAK1.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A. (ISOFORM 3).
RX MEDLINE=9820387; PubMed=9543386;
RA Peng C.Y., Graves P.R., O'G S., Thoma R.S., Byrnes M.J. III, Wu Z., Stephenson M.T., Pivnick-Worms H.;
RT "C-TAK1 protein kinase phosphorylates human Cdc25C on serine 216 and promotes 14-3-3 protein binding.";
RN Cell Growth Differ. 9:197-208(1998).
RN [2]
RP SEQUENCE FROM N.A. (ISOFORM 2).
RA Waggoner S.N., Stephen R., Farrar W.L., Howard O.M.Z.;
RT "Human serine/threonine protein kinase cTAK1/Kp78/Mark3: Identification of a novel splice variant and a larger 5'UTR.";
RN Submitted (JUN-1999) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A. (ISOFORM 4).
RX MEDLINE=21326471; PubMed=11433294;
RA Sun T.O., Lu B., Feng J.J., Reinhard C., Jan Y.N., Fantl W.J., Williams L.T.;
RT "PAR-1 is a Dishevelled-associated kinase and a positive regulator of Wnt signalling.";
RN Nat. Cell Biol. 3:628-636(2001).
RN [4]
RP SEQUENCE FROM N.A. (ISOFORM 5).
RA Drewes G.;
RT "Characterization of an alternatively spliced form of MARK3 from human brain.";
RN Submitted (JAN-2002) to the EMBL/GenBank/DBJ databases.
RN [5]
RP SEQUENCE FROM N.A. (ISOFORM 6).
RA Maheshwari K.K., Som S., Parsa I.;
RN Submitted (JAN-1992) to the EMBL/GenBank/DBJ databases.
RN [6]
RP SEQUENCE FROM N.A. (ISOFORM 3).

TISSUE=Cervix;
RC MEDLINE=22388257; PubMed=12477932;
RX Srausberg R.L., Feingold E.A., Grouse L.H., Derge J.G., Klausner R.D., Collins F.S., Wagner L., Shennen C.M., Schuler G.D., Altschul S.F., Jordan B., Buetow K.H., Schaefer C.F., Bhat N.K., Hopkins R.F., Zordan H., Moore T., Max S.I., Wang J., Hsien F., Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L., Spatlenko M., Soares M.B., Donald M.F., Casavant T.L., Scheetz T.E., Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C., Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaby S.J., Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H., Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W., Vallalao D.K., Munz D.M., Sodergren E.J., Lu X., Gibbs R.A., Fahy J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A., Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G., Blakeley R.W., Touchman J.W., Green E.D., Dickson M.C., Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.N., Krzywinski M.I., Skalek U., Smailus D.E., Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [7]
RP SEQUENCE OF 323-775 FROM N.A. (ISOFORM 7).
RC TISSUE=Urinary bladder;
RA Reynolds C.H., Patel U.A., Anderton B.H.;
RT "Homo sapiens mRNA partial sequence for a protein kinase, STK10, similar to p78/C-TAK1.";
RL Submitted (JUL-1999) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: Involved in the specific phosphorylation of microtubule-associated proteins for tau, MAP2 and MAP4. Phosphorylated Cdc25C on Ser-216.
CC -!- CATALYTIC ACTIVITY: ATP + a protein = ADP + a phosphoprotein.
CC -!- ALTERNATIVE PRODUCTS:
CC -!- Event-Alternative splicing; Named isoforms=7;
CC Names=1;
CC IsoId=P27448-1; Sequence=Displayed;
CC Names=2; Synonyms=CTAK75a;
CC IsoId=P27448-2; Sequence=VSP_004944;
CC Names=3;
CC IsoId=P27448-3; Sequence=VSP_004942, VSP_004944;
CC Names=4;
CC IsoId=P27448-4; Sequence=VSP_004942, VSP_004945;
CC Names=5;
CC IsoId=P27448-5; Sequence=VSP_004942;
CC Names=6; Synonyms=p58;
CC IsoId=P27448-6; Sequence=VSP_004942, VSP_004943, VSP_004944;
CC Names=7;
CC IsoId=P27448-7; Sequence=VSP_004943;
CC -!- TISSUE SPECIFICITY: Ubiquitous.
CC -!- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES. MARK SUBFAMILY.
CC -!- SIMILARITY: Contains 1 kinase-associated (KAI) domain.
CC -!- SIMILARITY: Contains 1 UBA domain.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/annouce/> or send an email to license@isb-sib.ch).
CC -----
CC EMBL; U64205; AAC15093.1; -;
DR EMBL; AF159295; AAD48007.1; -;
DR EMBL; AF387637; AAK82367.1; -;
DR EMBL; AF465413; AAL69982.1; -;
DR EMBL; M80359; AAA59991.1; -;
DR EMBL; BC024773; AAH24773.1; -;
DR EMBL; AF170723; AAD51631.1; -;
DR HSP; Q63450; IAO6.
DR Genew; HGNC:6897; MARK3.
DR MIM; 602678; -;

```

DR InterPro; IPR001172; Kinase Cterm.
DR InterPro; IPR000719; Prot_Kinase.
DR InterPro; IPR002290; Ser Thr_kinase.
DR InterPro; IPR001245; Tyr_kinase.
DR InterPro; IPR000449; UBA_domain.
DR Pfam; PF02149; KAI; 1.
DR Pfam; PF00069; kinase; 1.
DR Pfam; PF00627; UBA; 1.
DR PRINTS; PR00109; TYRKINASE.
DR ProDom; PD000001; Prot_Kinase; 1.
DR SMART; SM00220; S_TK; 1.
DR SMART; SM00165; UBA; 1.
DR PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
DR PROSITE; PS00011; PROTEIN_KINASE_DOM; 1.
DR PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
DR PROSITE; PS00030; UBA; 1.
KW Transferase; Serine/threonine-protein kinase; ATP-binding;
KW Alternative splicing.
FT DOMAIN 56 330 PROTEIN KINASE.
FT DOMAIN 326 365 UBA.
FT NP_BIND 62 70 KAI.
FT BINDING 62 70 ATP (BY SIMILARITY).
FT ACT_SITE 85 85 ATP (BY SIMILARITY).
FT ACT_SITE 201 201 BY SIMILARITY.
FT VARSPPLIC 162 184 Missing (in isoform 3, isoform 4, isoform 5 and isoform 6).
FT VARSPPLIC 394 409 /FTID=VSP_004942.
FT VARSPPLIC 638 661 Missing (in isoform 6 and isoform 7).
FT VARSPPLIC 638 661 /FTID=VSP_004943.
FT VARSPPLIC 638 661 Missing (in isoform 2, isoform 3 and isoform 6).
FT VARSPPLIC 638 661 /FTID=VSP_004944.
FT CONFLICT 125 125 Missing (in isoform 4).
FT CONFLICT 139 139 E -> Q (IN REF. 5).
FT CONFLICT 149 149 E -> K (IN REF. 4 AND 5).
FT CONFLICT 448 448 R -> K (IN REF. 2).
FT CONFLICT 466 466 A -> G (IN REF. 5).
FT CONFLICT 479 479 S -> T (IN REF. 5).
FT CONFLICT 539 539 A -> D (IN REF. 5).
FT CONFLICT 626 626 N -> T (IN REF. 5).
FT CONFLICT 668 668 E -> K (IN REF. 5).
SQ SEQUENCE 776 AA; 86944 MW; A45496849070098 CRC64;

Query Match 20.0%; Score 637; DB 1; Length 776;
Best Local Similarity 28.0%; Pred. No. 9,2e-26;
Matches 196; Conservative 102; Mismatches 251; Indels 150; Gaps 20;

Qy 2 KVEREIALKLEIHPHVLKLDVYENKYLVLVLEHVSQGGELFDYLVKKGLTPEARKF 61
Db 99 KLFREVRIMKILNHPNIVKLPFEVETETKTLVIMYASGGVFDYLVAGHGRMKEARSK 158
Qy 62 FRQ-----IISALDFCHSHSICHRLDKPENLLDKNNIRIA 98
Db 159 FRQCGAQGTIKVQVDFLLSMTFFIVSAVQYCHQKRVIRDLKAENLLDADMNKIA 218
Qy 99 DFGMASLQVGSLLTSCGSPHYACPEVIRGEKDYGRKADYVSGVILFALLVGLPFDD 158
Db 219 DFGFSNEFTVGGKLDTFGCSPPYAPAEFLQCKYDGPEDVWSVGLVILTVLSGLPFDG 278
Qy 159 DNLQLEKLVKGVGHMPHFIPDQCQSLLRGMIEVDAAARRUTLBIHQHVIYGGKNEPE 218
Db 279 QNLKELREVLGRKYRIPFYMTDCENLKRLFLNPKIKRGTLQEQIMKDRWINAGHEEDE 338
Qy 219 PEQIPKRVQIRLPSLEDIDPDVLDMSHSIGCFDRNKLQDLLEENQEKMIYELL 278
Db 339 LKPFVE-----PELDSQKQKIDIMVGGYSGEE---IQESLSKKWYDEITATYLL 387
Qy 279 DRKERYPSQED-EDLPPEINEIDPPKRVYDSPLNRRHGKRRPERKSMELSVTD----- 330
Db 388 GRKS---SELDASDSSSSSLAKVRPSSDLNNTSG-QSPHHKVRQSVSSSQKRRYSD 443

```

```

Qy 331 -GGSPVPARRAI-----EMAQHGQRSRISGASSGLSTSPSSPRVTPHPSPRG 378
Db 444 HAGPAIPSVVAYPKRSQTSSTADSDLKDGISRRKSSGSAVGKGIAAPSPMLGNASPNK 503
Qy 379 SPLPTPKGTPVHTPKESPA-----GTPNTPPS----- 406
Db 504 ADIPERKKSST-VPSNTASGGMTRRNYVCSERTTADRHVSIQNGKENSTIPDQRTFVA 562
Qy 407 ---SPSVGGVPMRRL---NSIKNSFLGSPFRHRRKLQVTPPEMSNLTPESSPELAKKS 460
Db 563 STHSISSAATPDRIRPRGRTASRSTFGQPR-ERRATYNGPPASPSLSHEATPLSQTRS 621
Qy 461 -----WFGNFIS-LEKEEQIFVWIKKPLSSIKADIVHAFSLPSLSHVSQTS----- 509
Db 622 RGSNTLFSKLTSLTRNMSFRFKLPTEYRNG-----RYEGSSRNVSAEQDENKE 675
Qy 510 -----FRAEYKATGG--PAVFQKPVKFOVD-----ITYTEGEAQKEN--- 545
Db 676 AKPRSLRFTWSMKTTSMDPGDMREIRKVLDDANNCDYEQRRERFLFCVHGDGHAENLVQ 735
Qy 546 -----GIYSVTFTLLSGPSRRFRKRVETIQAO 573
Db 736 WEMEVCKLPRLSLNGVRFKEISGTSTAFKNIAKIANEL 774

RESULT 12
AAK2 RAT STANDARD; PRT; 552 AA.
AC Q09137;
DT 01-FEB-1995 (Rel. 31, Created)
DT 01-FEB-1995 (Rel. 31, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE 5'-AMP-activated protein kinase, catalytic alpha-2 chain (EC 2.7.1.1-)
DE (AMPK alpha-2 chain).
GN PKAA2 OR AMPK2 OR AMPK.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=101116;
RN [1].
RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
RC TISSUE=Liver;
RX MEDLINE=94209324; PubMed=7908907;
RA Carling D., Aguan K., Woods A., Verhoeven A.J.M., Beri R.K.,
RA Brennan C.H., Sidebottom C., Davison M.D., Scott J.;
RT "Mammalian AMP-activated protein kinase is homologous to yeast and
RT plant protein kinases involved in the regulation of carbon
RT metabolism.";
RL J. Biol. Chem. 269:11442-11448(1994).
RN [2].
RP SEQUENCE FROM N.A.
RC STRAIN=Sprague-Dawley; TISSUE=Liver;
RX MEDLINE=95234757; PubMed=7718624;
RA Gao G., Widner J., Stapleton D., Teh T., Cox T., Kemp B.E.,
RA Watters L.A.;
RT "Catalytic subunits of the porcine and rat 5'-AMP-activated protein
RT kinase are members of the SNF1 protein kinase family.";
RL Biochim. Biophys. Acta 1266:73-82(1995).
CC -1- FUNCTION: RESPONSIBLE FOR THE REGULATION OF FATTY ACID SYNTHESIS
CC BY PHOSPHORYLATION OF ACETYL-COA CARBOXYLASE. IT ALSO REGULATES
CC CHOLESTEROL SYNTHESIS VIA PHOSPHORYLATION AND INACTIVATION OF
CC HORMONE-SENSITIVE LIPIASE AND HYDROXYMETHYLGLUTARYL-COA REDUCTASE.
CC APPEARS TO ACT AS A METABOLIC STRESS-SENSING PROTEIN KINASE
CC SWITCHING OFF BIOSYNTHETIC PATHWAYS WHEN CELLULAR ATP LEVELS ARE
CC DEPLETED AND WHEN 5'-AMP RISES IN RESPONSE TO FUEL LIMITATION
CC AND/OR HYPOXIA.. THIS IS A CATALYTIC SUBUNIT.
CC -1- SUBUNIT: HETEROTRIMER OF A CATALYTIC SUBUNIT, A BETA AND A GAMMA
CC NON-CATALYTIC SUBUNITS.
CC -1- Event=Alternative splicing; Named isoforms=2;
CC Name=Long;
CC IsoId=Q09137-1; Sequence=Displayed;
CC Name=Short;

```

CC isoid=009137-2; Sequence=VSP 004949, VSP 004950;
CC Note=Lacks the sequence parts essential for kinase activity and
CC is therefore inactive;
CC -1- TISSUE SPECIFICITY: SKELETAL MUSCLE, LOWER LEVELS IN LIVER, HEART,
CC AND KIDNEY.
CC -1- INDUCTION: BY AMP.
CC -1- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES. SNF1
CC SUBFAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; Z29486; CAA82620.1; -;
CC EMBL; U12149; AAB85033.1; -;
CC PIR; A53621; A53621.
CC HSSP; Q63450; 1A06.
CC InterPro; IPR000719; Prot_kinase.
CC InterPro; IPR002290; Ser_thr_kinase.
CC InterPro; IPR001245; Tyr_pkinase.
CC Pfam; PF00069; pkinase; 1.
CC PRINTS; PR00109; TYRKINASE.
CC ProDom; PD000001; Prot_kinase; 1.
CC SMART; SM00220; S_TKc; 1.
CC PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
CC PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
CC PROSITE; PS00111; PROTEIN_KINASE_DOM; 1.
CC Transfaser; Serine/threonine-protein kinase; Fatty acid biosynthesis;
CC Phosphorylation; ATP-binding; Multigene family; Alternative splicing.
CC DOMAIN 16 268 PROTEIN KINASE.
CC NP_BIND 22 30 ATP (BY SIMILARITY).
CC BINDING 45 45 ATP (BY SIMILARITY).
CC ACT_SITE 139 139 BY SIMILARITY.
CC MOD_RES 172 172 PHOSPHORYLATION (AUTO-) (POTENTIAL).
CC VARSPLOC 32 388 Missing (in isoform Short).
CC VARSPLOC 392 552 /FTId=VSP 004949.
CC CONFLICT 355 355 /FTId=VSP 004950.
CC CONFLICT 462 462 M -> S (IN REF. 2).
CC SEQUENCE 552 AA; 62257 MW; 2829E07F674D89B1 CRC64;
Query Match 19.9%; Score 636; DB 1; Length 552;
Best Local Similarity 28.4%; Pred. No. 7e-26;
Matches 174; Conservative 95; Mismatches 202; Indels 142; Gaps 21;
QY 2 KVEREIALKLEHPHVLKLDVYENKKYLVLEHVSGBELFDYLVKGRITPKARKF 61
Db 60 KIKREIQNLKFRHPIIKLVQVISTPTDFWMEYVSGELFDYCKHGRVEVEARLL 119
QY 62 FQIISALDFCHSHICHRDLKPNELLDEKNNIRIADFGWASLQVGSLETSCGSPHY 121
Db 120 FQOILSAVDYCHRHNVHVRDLKPNVLLDAQWNAKIADFGLSNMSDGEFLRTSCGSPNY 179
QY 122 ACPEVIRGEKDGKADYVSCGVILFALLVGLALPDNDNLKLEKXKRGVFMHPHETPP 181
Db 180 APEVISGRLVAGPBDVILWSCGVILYALLCGLTFPDDEHVTFLKIRGGVFIPEYLNLR 239
QY 182 DCQSLRGMIEYDAARRLTLEHIOKHVIYIGGKNEPEPEQIPRKVQIRLSFLSDIDPD 241
Db 240 STATLHMLQVDPKRAIKIDIREHWF-----KQDLPSYL-----FPEDPSYDAN 286
QY 242 VLDS-----MHSICGFRDRNKLQDLIS-EEENQEKMIYFLLLDRKERYSQDE---- 290
Db 287 VIDDEAVKEVCEKFC--TESEWNLSYSGDPQDQLAVAYHLIIDNR-RIMNQASEFYLA 343
QY 291 DLPPRNEIDPPRKVDSPMLNRHGRKRPKSMVLSVTDGSGPVPARRATEMAHQGQRS 350
Db 344 SSPPTG-----SFWD-----DMAMH----- 358

QY 351 RSISGASSGLSTSPSSPRVTPHPSPRGSP--PTPKGTPTVHTPKESPAGTNPPTPPSP 408
Db 359 -----IPGGLKHPPE-RMPFLADSPK-----RCPLDALNTTKPKSL 395
QY 409 SVGGVPRARLNSIKNSIFLGSFRFRHKLQVPTPEMSNL-----TPESPPELAKKSWFGN 464
Db 396 AVKAKAKWHLGIRSQSPDYDMAEVYRAMKQLDFEWKVVNAYHLRVRKQNPVT-----GN 449
QY 465 FISLEKEEQIFVVKDKPLSSIKADIVHAFSLSPSLSHSVISQTSFRAEYKATGGPAVFQ 524
Db 450 YVQMSL--QLYLV-----DNRSYLLDPKSIDDEVVEQSRSGSTPQRSCSAGLH 496
QY 525 KPVKFOVDITYTEGGBAQKENGIIYVTFLLSPSRFRFRVETIQALQLLSTDPDPAQAH 584
Db 497 RP-RSSVDSSTAENHS-----LSG-----SLTGLSLGTLSSASPRLGSH 535
QY 585 LSDTTN-CMEMMT 596
Db 536 TMDFFEMCASLIT 548
RESULT 13
AAK2 HUMAN STANDARD; PRT; 552 AA.
ID AAK2 HUMAN
AC P54646; Q9HLE8; Q9UD43;
DT 01-OCT-1996 (Rel. 34, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE 5'-AMP-activated protein kinase, catalytic alpha-2 chain (EC 2.7.1.1-)
DE (AMPK alpha-2 chain).
GN PRKAA2 OR AMPK2 OR AMPK.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Heart;
RX MEDLINE=95047501; PubMed=7959015;
RA Aguan K., Scott J., See C.G., Sarkar N.H.;
RT "Characterization and chromosomal localization of the human homologue
RT of a rat AMP-activated protein kinase-encoding gene: a major
RT regulator of lipid metabolism in mammals.";
RL Gene 149:345-350(1994).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Skeletal muscle;
RX MEDLINE=95080410; PubMed=7988703;
RA Beri R.K., Marley A.E., See C.G., Sopwith W.F., Aguan K., Carling D.,
RT Scott J., Carey F.;
RT "Molecular cloning, expression and chromosomal localisation of human
RT AMP-activated protein kinase.";
RL FEBS Lett. 356:117-121(1994).
RN [3]
RP SEQUENCE OF 33-552 FROM N.A.
RC Cobley V.;
RL Submitted (APR-2000) to the EMBL/GenBank/DBJ databases.
CC -1- FUNCTION: RESPONSIBLE FOR THE REGULATION OF FATTY ACID SYNTHESIS
CC BY PHOSPHORYLATION OF ACETYL-COA CARBOXYLASE. IT ALSO REGULATES
CC CHOLESTEROL SYNTHESIS VIA PHOSPHORYLATION AND INACTIVATION OF
CC HORMONE-SENSITIVE LIPASE AND HYDROXYMETHYLGUTARYL-COA REDUCTASE.
CC APPEARS TO ACT AS A METABOLIC STRESS-SENSING PROTEIN KINASE
CC SWITCHING OFF BIOSYNTHETIC PATHWAYS WHEN CELLULAR ATP LEVELS ARE
CC DEPLETED AND WHEN 5'-AMP RISES IN RESPONSE TO FUEL LIMITATION
CC AND/OR HYPOXIA. THIS IS A CATALYTIC SUBUNIT.
CC -1- SUBUNIT: HETEROTRIMER OF A CATALYTIC SUBUNIT, A BETA AND A GAMMA
CC NON-CATALYTIC SUBUNITS.
CC -1- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
CC SNF1 SUBFAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

```
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; U06454; AAA64745.1; -.
DR EMBL; AL035705; CAC17574.1; -.
DR PIR; S51025; S51025.
DR HSP; Q63450; 1A06.
DR Genew; HGNC:9377; PRKAA2.
DR MIM; 600497; -.
DR GO; GO:006468; P:protein amino acid phosphorylation; TAS.
DR InterPro; IPR000165; P:signal transduction; TAS.
DR InterPro; IPR000719; Prot_kinase.
DR InterPro; IPR002290; Ser_thr_kinase.
DR Pfam; PF00069; pkinase; 1.
DR ProDom; PD000001; Prot_kinase; 1.
DR SMART; SM00220; S_TK; 1.
DR PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
DR PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
DR PROSITE; PS00111; PROTEIN_KINASE_DOM; 1.
DR Transfaser; Serine/threonine-protein kinase; Fatty acid biosynthesis;
KW Phosphorylation; ATP-binding; Multigene family.
FT DOMAIN 16 268
FT NP_BIND 22 30 ATP (BY SIMILARITY).
FT BINDING 45 45 ATP (BY SIMILARITY).
FT ACT_SITE 139 139 BY SIMILARITY.
FT MOD_RES 172 172 PHOSPHORYLATION (AUTO-) (POTENTIAL).
FT CONFLICT 180 180 A -> T (IN REF. 1).
FT CONFLICT 271 271 D -> G (IN REF. 1).
FT CONFLICT 403 404 HL -> RO (IN REF. 1).
SQ SEQUENCE 552 AA; 62319 MW; C46AARCLD5104975 CRC64;

Query Match 19.8%; Score 632; DB 1; Length 552;
Best Local Similarity 27.9%; Pred. No. 1.1e-25;
Matches 171; Conservative 94; Mismatches 206; Indels 142; Gaps 20;

Qy 2 KVEREIALKLEHPLVHLKLDVYENKYLVLVLEHVGSGELFDYLVKKGLRTPKEARKP 61
Db :|||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
60 KIKREIQNLKFRHPIIKLYQVISTPTDFFWMEYVSGSGLFDYCKHGRVEMEARRL 119
Qy :|||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
62 FRQIISALDFCHSHSICHRDLKPNLLLDKNNIRIADFGMSLQVDSLLETSCGSPHY 121
Db :|||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
120 FQQLSALVDYCHRMVHRDLKPNVLDHAKNAIADFGLSNMWSDGFRTSCGSPNY 179
Qy :|||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
122 ACPEVIRGEKYDKRADVWCGVILFALLGALPPDDNLRLQLLEKVKRGVFMHPHFIPP 181
Db :|||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
180 AAPEVISGRLYAGPEVDIWSGCVILYALLCGTLPPDDEHVPTLFKIRGGVFYIPEYLN 239
Qy :|||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
182 DCQSLLRGMIEVDARRLTLEHQHVIWVGKNEPEPEQIPRKVQIRSLPSLEDIDPD 241
Db :|||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
240 SVATLLMMLQVDPDKRATIKDIREHEWF-----KQDLPSYL----FPDPSYDAN 286
Qy :|||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
242 VLDS-----MHSGLCFDRNKLQDLAS--BEENQKMIYFLLDRKERYPSQDE---- 290
Db :|||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
287 VIDDEAVKEVCEKEFC--TESEVMNLSYSGPQDLAVAYHLIDNR-RIMNQSEFYLA 343
Qy :|||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
291 DLPPRNEIDPPKRKRVDSPLNLRHGRKRRPERKSMELSVLTDGSPVPARRAJEMAHGQRS 350
Db :|||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
344 SSPPSGSF-----MDDSAMHIP----- 360
Qy :|||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
351 RSISGASGLSTSLSSPRVTPHPSRGSPL--PTPKGTPTVHTPKESPAGTPTPTSSP 408
Db :|||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
361 -----PGLKPHPE-RMPPLIADSPKA-----RCPLDALNTTKPKSL 395
Qy :|||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
409 SVGGVPMRRLNLSIKNSFLGSPRFRHRLQVTPTEMSNL-----TPSSPELAKKSWFGN 464
Db :|||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
396 AVKAKAKHLGIRSQSKPDIAMAEVTRAMKQLDFEKKVNVAYHLRVRKXNPT-----GN 449
Qy :|||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
465 FLSLEKEQIFVVIKDKPLSSIKADIVHAFISIPSLSHSVISQTSFRAEYKATGGPAVFQ 524
Db :|||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
450 YKGNLS--QLYLV-----DNRSYLLDFKSIDDEWVEQRSGSSTPQRSASAGLH 496
```

```
Qy 525 KPVKQVDITYTEGGEAQKENGNIYVTFITLLSGPSRRFRKRVETIQALLSTHDPAAQH 584
Db :|||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
497 RP-RSSFDSTAESH-----LSG-----SLTGLSTGLTSLSSVPRLGSH 535
Qy 585 LSDTTN-CMEMMT 596
Db :|||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
536 TMDEFFENCASLIIT 548

RESULT 14
AAKL RAT STANDARD; PRT; 548 AA.
AC P54645, 1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE 5'-AMP-activated protein kinase, catalytic alpha-1 chain (EC 2.7.1.-)
DE (AMPK alpha-1 chain).
GN PRKAA1 OR AMPK1.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
RP STRAIN=Sprague-Dawley; TISSUE=Hypothalamus, and Liver;
RX MEDLINE=96132781; PubMed=8557660;
RA Stapleton D., Mitchell K.I., Gao G., Widmer J., Michell B.J.,
RA Teh T., House C.M., Fernandez C.S., Cox T., Witters L.A.,
RA Kemp B.E.;
RT "Mammalian AMP-activated protein kinase subfamily.";
RL J. Biol. Chem. 271:611-614 (1996).
[2]
PARTIAL SEQUENCE.
RP STRAIN=Sprague-Dawley; TISSUE=Hypothalamus, and Liver;
RX MEDLINE=95050763; PubMed=7961907;
RA Stapleton D., Gao G., Michell B.J., Widmer J., Mitchell K.,
RA Teh T., House C.M., Witters L.A., Kemp B.E.;
RT "Mammalian 5'-AMP-activated protein kinase non-catalytic subunits are
RT homologs of proteins that interact with yeast Snf1 protein kinase.";
RL J. Biol. Chem. 269:29343-29346 (1994).
CC -1- FUNCTION: RESPONSIBLE FOR THE REGULATION OF FATTY ACID SYNTHESIS
CC BY PHOSPHORYLATION OF ACETYL-COA CARBOXYLASE. IT ALSO REGULATES
CC CHOLESTEROL SYNTHESIS VIA PHOSPHORYLATION AND INACTIVATION OF
CC HORMONE-SENSITIVE LIPASE AND HYDROXYMETHYLGUTARYL-COA REDUCTASE.
CC APPEARS TO ACT AS A METABOLIC STRESS-SENSING PROTEIN KINASE
CC SWITCHING OFF BIOSYNTHETIC PATHWAYS WHEN CELLULAR ATP LEVELS ARE
CC DEPLETED AND WHEN 5'-AMP RISES IN RESPONSE TO FUEL LIMITATION
CC AND/OR HYPOXIA. THIS IS A CATALYTIC SUBUNIT.
CC -1- SUBUNIT: HETEROTRIMER OF A CATALYTIC SUBUNIT, A BETA AND A GAMMA
CC NON-CATALYTIC SUBUNITS.
CC -1- TISSUE SPECIFICITY: LOW EXPRESSION IN KIDNEY, LIVER, LUNG, HEART,
CC AND BRAIN.
CC -1- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
CC SNF1 SUBFAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
```

```
CC EMBL; U40819; AAC52355.1; -.
DR HSP; Q63450; 1A06.
DR GO; GO:0005622; C:intracellular; ISS.
DR GO; GO:0008389; F:apoptosis inhibitor activity; ISS.
DR GO; GO:0004691; F:CAMP-dependent protein kinase activity; ISS.
DR GO; GO:0042557; F:eukaryotic elongation factor-2 kinase activ. .; NAS.
DR GO; GO:0000187; P:activation of MAPK; ISS.
DR GO; GO:0046318; P:negative regulation of glucosylceramide bio. .; ISS.
```

DR GO: 0017148; P: negative regulation of protein biosynthesis; NAS.
DR GO: 0045768; P: positive regulation of anti-apoptosis; ISS.
DR GO: 0045542; P: positive regulation of cholesterol biosynthesis; ISS.
DR GO: 0046321; P: positive regulation of fatty acid oxidation; NAS.
DR GO: 0046722; P: positive regulation of gluconeogenesis; NAS.
DR GO: 0046326; P: positive regulation of glucose import; NAS.
DR GO: 0006468; P: protein amino acid phosphorylation; NAS.
DR GO: 0001666; P: response to hypoxia; ISS.
DR InterPro: IPR000719; Prot_kinase.
DR InterPro: IPR002290; Ser_thr_kinase.
DR InterPro: IPR01245; Tyr_kinase.
DR Pfam: PF00069; pkinase; 1.
DR PRINTS: PR00109; TYRKINASE.
DR ProDom: PD000001; Prot_kinase; 1.
DR SMART: SM0220; S_TKc; 1.
DR PROSITE: PS00107; PROTEIN KINASE ATP; 1.
DR PROSITE: PS00108; PROTEIN KINASE ST; 1.
DR PROSITE: PS0011; PROTEIN KINASE DOM; 1.
KW Transferase; Serine/threonine-protein kinase; Fatty acid biosynthesis;
KW phosphorylation; ATP-binding; Multigene family.
FT DOMAIN 16 268 PROTEIN KINASE
FT NP_BIND 22 30 ATP (BY SIMILARITY)
FT BINDING 45 45 ATP (BY SIMILARITY)
FT ACT_SITE 139 139 BY SIMILARITY
FT MOD_RES 172 172 PHOSPHORYLATION (AUTO-) (POTENTIAL).
SQ SEQUENCE 548 AA; 62599 MW; 5CCAC3281C195F867 CRC64;

Query Match 19.8%; Score 631.5; DB 1; Length 548;
Best Local Similarity 40.7%; Pred. No. 1.2e-25;
Matches 134; Conservative 57; Mismatches 103; Indels 35; Gaps 8;

Qy 2 KVEREIALKLEIHPHVLKHDVYENKYLVLVLEHVGSGELFDYLVKGRGLTPKEARKF 61
Db 60 KIRREIQNLKFRPHIILKQVISTPDSDFWVEYVSGGELFDYICKNGELDEKERRL 119
Qy 62 FROIISALDFCHSHSICHRDLKPNLLDKENIRIADFGMASIQVGSLLSTSCGSPHY 121
Db 120 FQQLISGLVDYCHRWVHVRDLKPNVLLDAHNAKIADFGLSNMMSDGEFLRTSCGSPNY 179
Qy 122 ACPEVIRGEKVDGRKADVWCGVILFALLVGLPDDDLNLRQLLEKVKRGVFMHPHIPP 181
Db 180 AAPEVIRGLVAGPEVDVWSSGVILYALLCTGLPDDDDHVTPLFKKICDGFYTPQVILNP 239
Qy 182 DCQSLLRGMEIVDAARLLTLEHIQKHIWYIGGKNEPEPEQIPRKVQIRSLPSLED--ID 239
Db 240 SVISILKHLQVDPKRAIKDIREHEWF-----KQDLP-KYLPPEPSYSTWID 289
Qy 240 PDVLDSM-----HSLGCFRDRNKLQDLLEENQEKMIYFLLLDRKERYPSQDE 290
Db 290 DEALKEVCEKFECSSEEVLSCLYRNH--QDPLA-----VAYHLIIDNRIMNEAKDF 340
Qy 291 DLPRNEIDPPRKRVDSPLNAR-HGKRPP 318
Db 341 YL-----ATSPDPSFLDDHHLTRPHERVP 365

RESULT 15
SNIL_MOUSE
ID SNIL_MOUSE STANDARD; PRT; 779 AA.
AC Q60670;
DT 01-NOV-1997 (Rel. 35, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Probable serine/threonine protein kinase SNFLK (EC 2.7.1.-) (HRT-20)
DE (Myocardial SNF1-like kinase).
GN SNFLK OR MSK.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
CX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A., AND REVISIONS TO 8; 16-18; 44-45; 316 AND 435.
RC TISSUE=Embryo;

RA Ruiz J.C.;
RL Submitted (OCT-1999) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE OF 1-435 FROM N.A., AND TISSUE SPECIFICITY.
RC TISSUE=Embryo;
RX MEDLINE=95200798; PubMed=7893599;
RA Ruiz J.C., Conlon F.L., Robertson E.J.;
RT "Identification of novel protein kinases expressed in the myocardium of the developing mouse heart."
RL Mech. Dev. 48:153-164(1994).
CC -1- TISSUE SPECIFICITY: EXPRESSED IN LUNG, SKIN, OVARY, HEART AND STOMACH. NO EXPRESSION IN BRAIN, LIVER OR SKELETAL MUSCLE.
CC -1- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES. SNF1 SUBFAMILY.
CC -1- SIMILARITY: Contains 1 UBA domain.
CC
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).
CC
CC EMBL; U11494; AAA67926.2; -.
DR HSP; P24941; 1A01.
DR MGD; MGI:104754; Snflk.
DR InterPro: IPR000719; Prot_kinase.
DR InterPro: IPR002290; Ser_thr_kinase.
DR InterPro: IPR000449; UBA_domain.
DR Pfam: PF00069; pkinase; 1.
DR ProDom: PD000001; Prot_kinase; 1.
DR SMART: SM0220; S_TKc; 1.
DR PROSITE: PS00107; PROTEIN KINASE ATP; 1.
DR PROSITE: PS0011; PROTEIN KINASE DOM; 1.
DR PROSITE: PS00108; PROTEIN KINASE ST; 1.
DR PROSITE: PS00030; UBA; 1.
KW Transferase; Serine/threonine-protein kinase; ATP-binding.
FT DOMAIN 27 278 PROTEIN KINASE.
FT DOMAIN 303 343 UBA.
FT NP_BIND 33 41 ATP (BY SIMILARITY).
FT BINDING 56 56 ATP (BY SIMILARITY).
FT ACT_SITE 149 149 BY SIMILARITY.
SQ SEQUENCE 779 AA; 85027 MW; 7B08131BC46D9C4E CRC64;

Query Match 19.8%; Score 631; DB 1; Length 779;
Best Local Similarity 32.7%; Pred. No. 1.9e-25;
Matches 163; Conservative 74; Mismatches 170; Indels 92; Gaps 14;

Qy 2 KVEREIALKLEIHPHVLKHDVYENKYLVLVLEHVGSGELFDYLVKGRGLTPKEARKF 61
Db 70 KIYEVQLMKLLNHPNIIKLYQVWETKMLYIVTEFAQNGEMFDYLTNGLHSEARQK 129
Qy 62 FROIISALDFCHSHSICHRDLKPNLLDKENIRIADFGMASIQVGSLLSTSCGSPHY 121
Db 130 FWOILSAVEYCHNHVHRDLKTENLLDSDNMDIKLADFGFNGFYPKEPELSTCVGSPY 189
Qy 122 ACPEVIRGEKVDGRKADVWCGVILFALLVGLPDDDLNLRQLLEKVKRGVFMHPHIPP 181
Db 190 AAPEVIRGEKVDGRKADVWCGVILFALLVGLPDDDLNLRQLLEKVKRGVFMHPHIPP 249
Qy 182 DCQSLLRGMEIVDAARLLTLEHIQKHIWYIGGKNEPEPEQIPRKVQIRSLPSLEDI 238
Db 250 DCETLIRMLVDPKRAITTAIRQHRW-----QADPTLQDDDPADFMQGVTSNLGDY 304
Qy 239 DPDLVDSMHSILGCFRDRNKLQDLLEENQEKMIYFLLLDRKERY-----PSQDEDLPP 294
Db 305 NEQVLGIMQALGI--DRQRTIESLQNSYNHFAIYLLERLKEHRSQPSRPTAPT 362
Qy 295 R-----NEI--DPPRKRVDSM-----LNHRGKRPP--- 318
Db 363 RQQLRSSDLSLEVPOEILPCDPFRPSLLCPQALAQSVLQAEIDCDLHSLQLPLFP 422

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 26, 2003, 11:51:21 / Search time 32.4013 Seconds
(without alignments)
4842.281 Million cell updates/sec

Title: US-10-054-579-4
Perfect score: 3188
Sequence: 1 MKVERETAILKLEHPPHVLK.....TNCMEMTGRSLKCGIIPKS 608

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 830525 seqs, 258052604 residues

Total number of hits satisfying chosen parameters: 830525

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SPTREMBL 23.3*

- 1: sp_archaea.*
- 2: sp_bacteria.*
- 3: sp_fungi.*
- 4: sp_human.*
- 5: sp_invertebrate.*
- 6: sp_mammal.*
- 7: sp_mhc.*
- 8: sp_organelle.*
- 9: sp_phage.*
- 10: sp_plant.*
- 11: sp_rodent.*
- 12: sp_virus.*
- 13: sp_vertebrate.*
- 14: sp_unclassified.*
- 15: sp_rvirus.*
- 16: sp_bacteriap.*
- 17: sp_archaeap.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	3182	99.8	668	Q81WQ3	Q81WQ3 homo sapien
2	3015	94.6	603	O60843	O60843 homo sapien
3	2545.5	79.8	715	Q96JL4	Q96JL4 homo sapien
4	2545.5	79.8	768	Q8ND00	Q8ND00 homo sapien
5	2545.5	79.8	778	Q8TDC2	Q8TDC2 homo sapien
6	2545.5	79.8	794	Q8TDC3	Q8TDC3 homo sapien
7	1712	53.7	793	Q9VUV4	Q9VUV4 drosophila
8	1596.5	50.1	698	O61298	O61298 halocynthia
9	1536	48.2	914	Q19469	Q19469 caenorhabdi
10	1439.5	45.2	473	Q96AV4	Q96AV4 homo sapien
11	1248	39.1	701	Q95T82	Q95T82 drosophila
12	769.5	24.1	833	Q8SXS5	Q8SXS5 dictyosteli
13	718	22.5	170	Q8TB60	Q8TB60 homo sapien
14	698	21.9	401	Q98SC8	Q98SC8 guillardia
15	698	21.9	535	Q8RW22	Q8RW22 arabidopsis
16	688.5	21.6	745	Q15524	Q15524 homo sapien

17	688.5	21.6	755	4	Q96HB3	Q96hb3 homo sapien
18	686.5	21.5	691	4	Q96RG0	Q96rg0 homo sapien
19	686	21.5	504	10	P93113	P93113 cucumis sat
20	684.5	21.5	722	11	O08679	O08679 rattus norv
21	683.5	21.4	888	11	Q8BR95	Q8br95 mus musculu
22	681.5	21.4	512	10	P92968	P92968 arabidopsis
23	681.5	21.4	512	10	P92958	P92958 arabidopsis
24	680	21.3	149	4	O95099	O95099 homo sapien
25	680	21.3	505	10	Q92RJ1	Q92rj1 cryza sativ
26	679	21.3	514	10	Q9XP25	Q9xf25 glycine max
27	676.5	21.2	785	13	Q8QGV3	Q8qgv3 xenopus lae
28	675.5	21.2	752	4	Q8NG37	Q8ng37 homo sapien
29	672.5	21.1	484	10	Q43475	Q43475 hordeum vul
30	672.5	21.1	926	4	Q9H0K1	Q9h0k1 homo sapien
31	671.5	21.1	729	11	Q9JKE4	Q9jke4 mus musculu
32	671.5	21.1	744	11	Q9JKE5	Q9jke5 mus musculu
33	670.5	21.0	795	4	Q9POL2	Q9pol2 homo sapien
34	670	21.0	514	10	Q9M726	Q9m726 lycopersico
35	668	21.0	511	10	Q40544	Q40544 nicotiana t
36	668	21.0	931	11	Q8CFH6	Q8cfh6 mus musculu
37	667.5	20.9	752	11	Q8CIP4	Q8cip4 mus musculu
38	665.5	20.9	671	3	Q96W17	Q96w17 trichoderma
39	663.5	20.8	626	5	Q95ZQ4	Q95zq4 caenorhabdi
40	660.5	20.7	624	5	Q22068	Q22068 caenorhabdi
41	658.5	20.7	797	11	Q8VHF0	Q8vhf0 rattus norv
42	657.5	20.6	793	11	O08678	O08678 rattus norv
43	657.5	20.6	795	11	Q8VHJ5	Q8vhj5 mus musculu
44	657.5	20.6	1371	4	Q9Y2K2	Q9y2k2 homo sapien
45	655	20.5	1037	3	Q8NKJ8	Q8nkj8 saccharomyc

ALIGNMENTS

RESULT 1

Q81WQ3 PRELIMINARY; PRT; 668 AA.
AC Q81WQ3: 01-MAR-2003 (TREMREL. 23, Created)
DT 01-MAR-2003 (TREMREL. 23, Last sequence update)
DT 01-MAR-2003 (TREMREL. 23, Last annotation update)
DE Serine/threonine protein kinase isoform.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
RX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Brain tumor;
RA Guo J.H., Yu L.;
RL Submitted (OCT-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY166857; AAN87839.1; -
KW Kinase.
SQ SEQUENCE 668 AA; 74714 MW; 6AF8CB84FC48C07 CRC64;

Query Match	99.8%;	Score 3182;	DB 4;	Length 668;
Best Local Similarity	99.8%;	Pred. No. 2.1e-229;		
Matches 607;	Conservative 0;	Mismatches 1;	Indels 0;	Gaps 0;
Qy	1	MKVERETAILKLEHPPHVLKLDVYENKYLVLVLEHVS	GGELFDYLVKKGRLTPKEARK	60
Db	61	MKVERETAILKLEHPPHVLKLDVYENKYLVLVLEHVS	GGELFDYLVKKGRLTPKEARK	120
Qy	61	FFRQIISALDFCHSHSICHRDLKPENLLDEKNIRIAD	FGMASLQVGDLSLETSCSPH	120
Db	121	FFRQIISALDFCHSHSICHRDLKPENLLDEKNIRIAD	FGMASLQVGDLSLETSCSPH	180
Qy	121	YACPEVIRGKYGDKADVWSCGVILFALLVGLALPDD	DLNRLQLLEKVGVFHMFIP	180
Db	181	YACPEVIRGKYGDKADVWSCGVILFALLVGLALPDD	DLNRLQLLEKVGVFHMFIP	240
Qy	181	PDQCSLRGMIEVDAAARLTLFHIQHIWYIGGKNEPE	PEPQIPRKVIQIRSLPSLEIDP	240

```
Db 241 PDCQSLRGMSVDAARELTLEHIQKIWIYGGKNEPEPEQIPRKQVIRSLPSLEIDP 300
Qy 241 DVLDSMSLGCGRDRNKLQDLSEENQEKIYIFLLDRKERYPSQDEDELPNRNIDP 300
Db 301 DVLDSMSLGCGRDRNKLQDLSEENQEKIYIFLLDRKERYPSQDEDELPNRNIDP 360
Qy 301 PRKRVDSPLNRHGKRRRPERKSMVLSVTDGSGSPVARRAIEAQAQHGORSISGASGL 360
Db 361 PRKRVDSPLNRHGKRRRPERKSMVLSVTDGSGSPVARRAIEAQAQHGORSISGASGL 420
Qy 361 STPSLSSPRVTPHSPRGSPPLTPKGTVPVHTPKSPAGTNPPTPPSPVSGVPMRARN 420
Db 421 STPSLSSPRVTPHSPRGSPPLTPKGTVPVHTPKSPAGTNPPTPPSPVSGVPMRARN 480
Qy 421 SIKNSFLGSPFRHRKLOVPTPEENSLNTPSSPELAKKSWFGNFI SLEKEEQIFVVIKD 480
Db 481 SIKNSFLGSPFRHRKLOVPTPEENSLNTPSSPELAKKSWFGNFI SLEKEEQIFVVIKD 540
Qy 481 KPLSSIKADIHVAFLSIPSLSHSVISQTSFRAEYKATGPAVFOKPVKFQVDITVTEGGE 540
Db 541 KPLSSIKADIHVAFLSIPSLSHSVISQTSFRAEYKATGPAVFOKPVKFQVDITVTEGGE 600
Qy 541 AQKENGIVSVTFTLLSGSPRRFRVETIQALLSTHDPAAQHLSDTTNCMMWTGRLS 600
Db 601 AQKENGIVSVTFTLLSGSPRRFRVETIQALLSTHDPAAQHLSDTTNCMMWTGRLS 660
Qy 601 KCGIIPKS 608
Db 661 KCGIIPKS 668

RESULT 2
O60843 PRELIMINARY; PRT; 603 AA.
ID O60843
AC O60843
DT 01-AUG-1998 (TrEMBLrel. 07, Created)
DT 01-AUG-1998 (TrEMBLrel. 07, Last sequence update)
DE Putative serine/threonine protein kinase (Fragment).
DE Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Brain;
RX MEDLINE=21064499; PubMed=11124703;
RA Stanchi F., Bertocco E., Toppo S., Dioguardi R., Simionati B.,
RA Cannata N., Zimbello R., Lanfranchi G., Valle G.;
RT "Characterization of 16 novel human genes showing high similarity to
RT yeast sequences";
RL Yeast 18:69-80(2001).
CC -!- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
DR EMBL; AJ006701; CAA07196.1; -.
DR Genbank; HGNC:11405; STK29.
DR InterPro; IPR000719; Prot kinase.
DR InterPro; IPR002290; Ser_thr_kinase.
DR Pfam; PF00069; pkinase; 1.
DR ProDom; PD000001; Prot kinase; 1.
DR PROSITE; PS00011; PROTEIN KINASE DOM; 1.
DR PROSITE; PS00108; PROTEIN KINASE ST; 1.
KW ATP-binding; Kinase; Serine/threonine-protein kinase; Transferase.
FT NON_TER 1
SQ SEQUENCE 603 AA; 67401 MW; B02C5D678F8EE96 CRC64;

Query Match 94.6%; Score 3015; DB 4; Length 603;
Best Local Similarity 99.7%; Pred. No. 5.5e-217;
Matches 574; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 12 LIEHPVLKLDVYENKYLVLVLEHVSGLFDVLVKKGRITPKARKFFQIIISALDF 71
Db 1 LIEHPVLKLDVYENKYLVLVLEHVSGLFDVLVKKGRITPKARKFFQIIISALDF 60
```

```
Qy 72 CHSHSICHRLDLPENLLDEKNNIRIADFGMASLOVGDLSLETSCGSPHYACPEVIRGEK 131
Db 61 CHSHSICHRLDLPENLLDEKNNIRIADFGMASLOVGDLSLETSCGSPHYACPEVIRGEK 120
Qy 132 YDGRKADVMSGCVILFALLVGLPFPDDNLRQLLEKVKRGVFMHPHFIPDQCQSLRGMI 191
Db 121 YDGRKADVMSGCVILFALLVGLPFPDDNLRQLLEKVKRGVFMHPHFIPDQCQSLRGMS 180
Qy 192 EVDAAARLTLEHIQKIWIYGGKNEPEPEQIPRKQVIRSLPSLEIDPDLVDSMSLGC 251
Db 181 EVDAAARLTLEHIQKIWIYGGKNEPEPEQIPRKQVIRSLPSLEIDPDLVDSMSLGC 240
Qy 252 FRDRNKLQDLSEENQEKIYIFLLDRKERYPSQDEDELPNRNIDPDRKRVDSPLMN 311
Db 241 FRDRNKLQDLSEENQEKIYIFLLDRKERYPSQDEDELPNRNIDPDRKRVDSPLMN 300
Qy 312 RHGKRPERKSMVLSVTDGSGSPVARRAIEAQAQHGORSISGASGLSTPSLSPRVT 371
Db 301 RHGKRPERKSMVLSVTDGSGSPVARRAIEAQAQHGORSISGASGLSTPSLSPRVT 360
Qy 372 PHSPRGSPPLTPKGTVPVHTPKSPAGTNPPTPPSPVSGVPMRARNISIKNSFLGSPR 431
Db 361 PHSPRGSPPLTPKGTVPVHTPKSPAGTNPPTPPSPVSGVPMRARNISIKNSFLGSPR 420
Qy 432 FHRKLOVPTPEENSLNTPSSPELAKKSWFGNFI SLEKEEQIFVVIKDPLSSIKADI 491
Db 421 FHRKLOVPTPEENSLNTPSSPELAKKSWFGNFI SLEKEEQIFVVIKDPLSSIKADI 480
Qy 492 HAFSLPSLSHVISQTSFRAEYKATGPAVFOKPVKFQVDITVTEGGEAQKENGIVSVT 551
Db 481 HAFSLPSLSHVISQTSFRAEYKATGPAVFOKPVKFQVDITVTEGGEAQKENGIVSVT 540
Qy 552 FTLLSGSPRRFRVETIQALLSTHDPAAQHLSD 587
Db 541 FTLLSGSPRRFRVETIQALLSTHDPAAQHLSE 576

RESULT 3
Q96JL4
ID Q96JL4 PRELIMINARY; PRT; 715 AA.
AC Q96JL4
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DE 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Hypothetical protein KIAA1811 (Fragment).
GN KIAA1811.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Brain;
RX MEDLINE=21245130; PubMed=11347906;
RA Nagase T., Nakayama M., Nakajima D., Kikuno R., Ohara O.;
RA "Prediction of the coding sequences of unidentified human genes. XX.
RT The complete sequences of 100 new cDNA clones from brain which code
RT for large proteins in vitro.";
RL DNA Res. 8:85-95(2001).
CC -!- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
DR EMBL; AB058714; BAB47440.1; -.
DR HSSP; P24941; 18UH.
DR InterPro; IPR000719; Prot kinase.
DR InterPro; IPR002290; Ser_thr_kinase.
DR Pfam; PF00069; pkinase; 1.
DR ProDom; PD000001; Prot kinase; 1.
DR SMART; SM00220; S_TKc; 1.
DR PROSITE; PS00011; PROTEIN KINASE DOM; 1.
DR PROSITE; PS00108; PROTEIN KINASE ST; 1.
KW Hypothetical protein; ATP-binding; Kinase;
KW Serine/threonine-protein kinase; Transferase.
```

```

FT  NON_TER  1 1
SQ  SEQUENCE  715 AA; 78499 MW; B90F6EE115C418A5 CRC64;

Query Match
Best Local Similarity 79.8%; Score 2545.5; DB 4; Length 715;
Matches 495; Conservative 51; Mismatches 44; Indels 65; Gaps 7;

QY  1 MKVEREAILKLIIEHPVHLKLDVYENKYLVLVLEHVSOGGELFDYLVKGRGLTPKEARK 60
Db  13 MKVEREAILKLIIEHPVHLKLDVYENKYLVLVLEHVSOGGELFDYLVKGRGLTPKEARK 72

QY  61 FFRQIISALDFCHSHSICHRLDKPENLLDEKNIRIADFGMASLOVGDLSLETSCGSPH 120
Db  73 FFRQIVSALDFCHSYCHRLDKPENLLDEKNIRIADFGMASLOVGDLSLETSCGSPH 132

QY  121 YACPEVIRGEKYDGRKADVWSCGVILFALLVGVLPDNDNLRLQLEKVKRGVFMHPHIP 180
Db  133 YACPEVIRGEKYDGRADWMSGVILFALLVGVLPDNDNLRLQLEKVKRGVFMHPHIP 192

QY  181 PDCQSLLRGMIEVDAAARLTLEHIQKHIWYIGGNKNEPEP-EQPIP-RKVQIRSLPSLEDI 238
Db  193 PDCQSLLRGMIEVEPEKRLSLQIQKHPWYLGKHEPDPCLEPAPGRRVAMRSLPSNGEL 252

QY  239 DPVLDLSMHSICGFRDRNKLQDLISEENQEKMIYFLLDRKERYPSQEDDLPPRNEI 298
Db  253 DPVLDLSMHSICGFRDRERLHRLSEENQEKMIYFLLDRKERYPSQEDDLPPRNDV 312

QY  299 DPPRKRVDSPLNLRHGRKRRPERKSMVLSVTD---GGSPVPARRAIEMAHQGRSRSISG 355
Db  313 DPPRKRVDSPLNLRHGRKRRPERKSMVLSITDAGGGSPVPTRRALEMAHQGRSRSVSG 372

QY  356 ASSGLSTSPSSPR-----VTPHP 374
Db  373 ASTGLSSPLSSPSRSPVFSPEPGAGDEARGGSGPTSKTQTLPSRPGGGAGGQPPPP 432

QY  375 SPRGSLPTPKG-----TPVHTPKESPAGTNPPTPSSP--SVGGVPWRARLNS 421
Db  433 SARSTPLPGPPGSPRSSGGTPLHSPHTPRASPTGTPTTTPPPSPGGVGGAAWRSLNS 492

QY  422 IKNSFLGSPRFRHRRKLVQPTPEEMSNLTPESSPELAKKSWFGNFIISLEKEEIQIFVVIKDK 481
Db  493 INNSFLGSPRFRHRRKLVQPTAEEMSLTPESSPELAKKSWFGNFIISLEKEEIQIFVVIKDK 552

QY  482 PLSSIKADIVHAFISLPSLSHVSISQTSFRAEYKATGGPAVKPQKVFQVDITYTEGGEA 541
Db  553 PLSSIKADIVHAFISLPSLSHVSISQTSFRAEYKASGGSPVFKPVRFOVDISSEGEPEP 612

QY  542 QKE-----NGIYSVTFTLLSGPSRRFRKRVVETIQALLSTHDPAPAAQLSDTTN 590
Db  613 SPRRDGSGGGGIYSVTFTLLSGPSRRFRKRVVETIQALLSTHDPQSVQALADEKN 667

```

RESULT 4

```

Q8ND00
ID  Q8ND00  PRELIMINARY; PRT; 768 AA.
AC  Q8ND00;
DT  01-OCT-2002 (TrEMBLrel. 22, Created)
DT  01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT  01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE  Hypothetical protein (Fragment).
GN  DKF2547E1613.
OS  Homo sapiens (Human).
OC  Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC  Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX  NCBI_TaxID=9606;
RN  [1]
RP  SEQUENCE FROM N.A.
RC  TISSUE=Brain;
RA  Koehrer K., Beyer A., Mewes H.W., Weil B., Wiemann S.;
RL  Submitted (JUL-2002) to the EMBL/GenBank/DBJ databases.
DR  EMBL; AL834275; CAD38950.1;
DR  InterPro; IPR000719; Prot_kinase.
DR  InterPro; IPR002290; Ser_thr_kinase.

```

RESULT 5

```

Q8TDC2
ID  Q8TDC2  PRELIMINARY; PRT; 778 AA.
AC  Q8TDC2;
DT  01-JUN-2002 (TrEMBLrel. 21, Created)
DT  01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT  01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE  Protein kinase-like protein.
OS  Homo sapiens (Human).
OC  Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC  Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

```

```

DR  InterPro; IPR001245; Tyr_kinase.
DR  InterPro; IPR000449; UBA domain.
DR  Pfam; PF0069; pkinase; 1.
DR  ProDom; PD000001; Prot_kinase; 1.
DR  SMART; SM00220; S_TKC; 1.
DR  SMART; SM00219; TyrKc; 1.
DR  PROSITE; PS00107; PROTEIN KINASE ATP; 1.
DR  PROSITE; PS00101; PROTEIN KINASE DOM; 1.
DR  PROSITE; PS00108; PROTEIN KINASE ST; 1.
KW  Hypothetical protein; ATP-binding; Transferase.
FT  NON_TER  1
SQ  SEQUENCE  768 AA; 83962 MW; 1P954ADFF62B899A CRC64;

Query Match
Best Local Similarity 79.8%; Score 2545.5; DB 4; Length 768;
Matches 495; Conservative 51; Mismatches 44; Indels 65; Gaps 7;

QY  1 MKVEREAILKLIIEHPVHLKLDVYENKYLVLVLEHVSOGGELFDYLVKGRGLTPKEARK 60
Db  42 MKVEREAILKLIIEHPVHLKLDVYENKYLVLVLEHVSOGGELFDYLVKGRGLTPKEARK 101

QY  61 FFRQIISALDFCHSHSICHRLDKPENLLDEKNIRIADFGMASLOVGDLSLETSCGSPH 120
Db  102 FFRQIVSALDFCHSYCHRLDKPENLLDEKNIRIADFGMASLOVGDLSLETSCGSPH 161

QY  121 YACPEVIRGEKYDGRKADVWSCGVILFALLVGVLPDNDNLRLQLEKVKRGVFMHPHIP 180
Db  162 YACPEVIRGEKYDGRADWMSGVILFALLVGVLPDNDNLRLQLEKVKRGVFMHPHIP 221

QY  181 PDCQSLLRGMIEVDAAARLTLEHIQKHIWYIGGNKNEPEP-EQPIP-RKVQIRSLPSLEDI 238
Db  222 PDCQSLLRGMIEVEPEKRLSLQIQKHPWYLGKHEPDPCLEPAPGRRVAMRSLPSNGEL 281

QY  239 DPVLDLSMHSICGFRDRNKLQDLISEENQEKMIYFLLDRKERYPSQEDDLPPRNEI 298
Db  282 DPVLDLSMHSICGFRDRERLHRLSEENQEKMIYFLLDRKERYPSQEDDLPPRNDV 341

QY  299 DPPRKRVDSPLNLRHGRKRRPERKSMVLSVTD---GGSPVPARRAIEMAHQGRSRSISG 355
Db  342 DPPRKRVDSPLNLRHGRKRRPERKSMVLSITDAGGGSPVPTRRALEMAHQGRSRSVSG 401

QY  356 ASSGLSTSPSSPR-----VTPHP 374
Db  402 ASTGLSSPLSSPSRSPVFSPEPGAGDEARGGSGPTSKTQTLPSRPGGGAGGQPPPP 461

QY  375 SPRGSLPTPKG-----TPVHTPKESPAGTNPPTPSSP--SVGGVPWRARLNS 421
Db  462 SARSTPLPGPPGSPRSSGGTPLHSPHTPRASPTGTPTTTPPPSPGGVGGAAWRSLNS 521

QY  422 IKNSFLGSPRFRHRRKLVQPTPEEMSNLTPESSPELAKKSWFGNFIISLEKEEIQIFVVIKDK 481
Db  522 IRNSFLGSPRFRHRRKLVQPTAEEMSLTPESSPELAKKSWFGNFIISLEKEEIQIFVVIKDK 581

QY  482 PLSSIKADIVHAFISLPSLSHVSISQTSFRAEYKATGGPAVKPQKVFQVDITYTEGGEA 541
Db  582 PLSSIKADIVHAFISLPSLSHVSISQTSFRAEYKASGGSPVFKPVRFOVDISSEGEPEP 641

QY  542 QKE-----NGIYSVTFTLLSGPSRRFRKRVVETIQALLSTHDPAPAAQLSDTTN 590
Db  642 SPRRDGSGGGGIYSVTFTLLSGPSRRFRKRVVETIQALLSTHDPQSVQALADEKN 696

```


OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA She X.Y., Yu L., Guo J.H.;
RL Submitted (FEB-2002) to the EMBL/GenBank/DBJ databases.
CC -!- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
DR EMBL; AF479827; AAL87698.1; -
DR HSSP; P24941; 1BUH.
DR InterPro; IPR000719; Prot kinase.
DR InterPro; IPR002290; Ser_thr_kinase.
DR Pfam; PF00069; kinase; 1.
DR SMART; SM00220; S_TK; 1.
DR PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
DR PROSITE; PS00111; PROTEIN_KINASE_DOM; 1.
DR PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
KW ATP-binding; Kinase; Serine/threonine-protein kinase; Transferase.
SQ SEQUENCE 778 AA; 85086 MW; 8D1818D4E54398BB CRC64;

Query Match 79.8%; Score 2545.5; DB 4; Length 778;
Best Local Similarity 75.6%; Pred. No. 9.9e-182;
Matches 495; Conservative 51; Mismatches 44; Indels 65; Gaps 7;

QY 1 MKVERETAILKLIHHPVLKLDHVVYENKKYLYLVLEHVSQGLFDYLVKKGRITPKEAR 60
DB 76 MKVERETAILKLIHHPVLKLDHVVYENKKYLYLVLEHVSQGLFDYLVKKGRITPKEAR 135

QY 61 FFRQIISALDFCHSHSICHRLDKPENLLLDKNNIRIADFGMASLQVDSLLTSCGSPH 120
DB 136 FFRQIVSALDFCHSHSICHRLDKPENLLLDKNNIRIADFGMASLQVDSLLTSCGSPH 195

QY 121 YACPEVIRGEKYDGRADVNSCGVILFALLVGLALPDNDNLRLQLLEKVKRGVFMHPHFIP 180
DB 196 YACPEVIRGEKYDGRADVNSCGVILFALLVGLALPDNDNLRLQLLEKVKRGVFMHPHFIP 255

QY 181 PDCQSLLRGMIEVDAAARLTLEHIQKIHWYIGGKNEPEP-EQFIP-RKVQIRSLPSLEDI 238
DB 256 PDCQSLLRGMIEVPEKRLSLEQIQKHPWYLGKHEPDCLEPAPGRVAMRSLPSNGEL 315

QY 239 DPVLDMSHSLGCFDRNRKLLQDLSEENQEKMIYFLLLDKRYEYPSQDEDLPPRNEI 298
DB 316 DPVLESMSLGCGRDRERLHRELSEENQEKMIYFLLLDKRYEYPSQDEDLPPRNDV 375

QY 299 DPPRKRVDSPLNRHGRKRRPERKSMVLSVTD---GGSPVPARRAIEMAHQGRSRSISG 355
DB 376 DPPRKRVDSPLNRHGRKRRPERKSMVLSITDAGGGGSPVTRRALEMAHQGRSRSVSG 435

QY 356 ASSGLSTSPSSPR-----VTTPH 374
DB 436 ASTGLSSPLSSPRSPVFSFSPGAGDEARGGSGTSTQTTLPSRGRGGGAGEQPPPP 495

QY 375 SPRGSLPTPKG-----TPVHTPKESAGTNPPTPPSP--SVGGVPMRLNS 421
DB 496 SARSTPLPGPPGPRSSGGTPLHSPHLTPRASPTGTPTTTPPSPGGVGGAAWRSLNS 555

QY 422 IKNSFLGSPRFRHRRKLVQPTPEEMSNLTPESSPELAKSWFGNFIISLEKEQIFVWIKDK 481
DB 556 IRNSFLGSPRFRHRRKLVQPTPEEMSNLTPESSPELAKSWFGNFIISLEKEQIFVWIKDK 615

QY 482 PLSSIKADIVHAFSLIPSLSHSVISQTSFRAEYKATGGPAVQKPVKQVFDITYTEGEA 541
DB 616 PLSSIKADIVHAFSLIPSLSHSVISQTSFRAEYKATGGPAVQKPVKQVFDITYTEGEA 675

QY 542 QKE-----NGIYSVTFTLLSGSPRRFRKRVVETIQALLSTHDPAPAAQLSDTTN 590
DB 676 SPRDGGGGGGIYVFTLLSGSPRRFRKRVVETIQALLSTHDPAPAAQLADEKN 730

RESULT 6

Q8TDC3

ID Q8TDC3

AC Q8TDC3;

PRELIMINARY; PRT; 794 AA.

DT 01-JUN-2002 (TReMBLrel. 21, Created)
DT 01-JUN-2002 (TReMBLrel. 21, Last sequence update)
DT 01-MAR-2003 (TReMBLrel. 23, Last annotation update)
DE Putative serine/threonine protein kinase.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Brain;
RA She X.Y., Guo J.H., Yu L.;
RL Submitted (FEB-2002) to the EMBL/GenBank/DBJ databases.
CC -!- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
DR EMBL; AF479826; AAL87697.1; -
DR HSSP; P24941; 1BUH.
DR InterPro; IPR000719; Prot kinase.
DR InterPro; IPR002290; Ser_thr_kinase.
DR Pfam; PF00069; kinase; 1.
DR ProDom; PD000001; Prot kinase; 1.
DR SMART; SM00220; S_TK; 1.
DR PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
DR PROSITE; PS00111; PROTEIN_KINASE_DOM; 1.
DR PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
KW ATP-binding; Kinase; Serine/threonine-protein kinase; Transferase.
SQ SEQUENCE 794 AA; 86753 MW; 5DD395B0B61ABF77 CRC64;

Query Match 79.8%; Score 2545.5; DB 4; Length 794;
Best Local Similarity 75.6%; Pred. No. 1e-181;
Matches 495; Conservative 51; Mismatches 44; Indels 65; Gaps 7;

QY 1 MKVERETAILKLIHHPVLKLDHVVYENKKYLYLVLEHVSQGLFDYLVKKGRITPKEAR 60
DB 92 MKVERETAILKLIHHPVLKLDHVVYENKKYLYLVLEHVSQGLFDYLVKKGRITPKEAR 151

QY 61 FFRQIISALDFCHSHSICHRLDKPENLLLDKNNIRIADFGMASLQVDSLLTSCGSPH 120
DB 152 FFRQIVSALDFCHSHSICHRLDKPENLLLDKNNIRIADFGMASLQVDSLLTSCGSPH 211

QY 121 YACPEVIRGEKYDGRADVNSCGVILFALLVGLALPDNDNLRLQLLEKVKRGVFMHPHFIP 180
DB 212 YACPEVIRGEKYDGRADVNSCGVILFALLVGLALPDNDNLRLQLLEKVKRGVFMHPHFIP 271

QY 181 PDCQSLLRGMIEVDAAARLTLEHIQKIHWYIGGKNEPEP-EQFIP-RKVQIRSLPSLEDI 238
DB 272 PDCQSLLRGMIEVPEKRLSLEQIQKHPWYLGKHEPDCLEPAPGRVAMRSLPSNGEL 331

QY 239 DPVLDMSHSLGCFDRNRKLLQDLSEENQEKMIYFLLLDKRYEYPSQDEDLPPRNEI 298
DB 332 DPVLESMSLGCGRDRERLHRELSEENQEKMIYFLLLDKRYEYPSQDEDLPPRNDV 391

QY 299 DPPRKRVDSPLNRHGRKRRPERKSMVLSVTD---GGSPVPARRAIEMAHQGRSRSISG 355
DB 392 DPPRKRVDSPLNRHGRKRRPERKSMVLSITDAGGGGSPVTRRALEMAHQGRSRSVSG 451

QY 356 ASSGLSTSPSSPR-----VTTPH 374
DB 452 ASTGLSSPLSSPRSPVFSFSPGAGDEARGGSGTSTQTTLPSRGRGGGAGEQPPPP 511

QY 375 SPRGSLPTPKG-----TPVHTPKESAGTNPPTPPSP--SVGGVPMRLNS 421
DB 512 SARSTPLPGPPGPRSSGGTPLHSPHLTPRASPTGTPTTTPPSPGGVGGAAWRSLNS 571

QY 422 IKNSFLGSPRFRHRRKLVQPTPEEMSNLTPESSPELAKSWFGNFIISLEKEQIFVWIKDK 481
DB 572 IRNSFLGSPRFRHRRKLVQPTPEEMSNLTPESSPELAKSWFGNFIISLEKEQIFVWIKDK 631

QY 482 PLSSIKADIVHAFSLIPSLSHSVISQTSFRAEYKATGGPAVQKPVKQVFDITYTEGEA 541
DB 632 PLSSIKADIVHAFSLIPSLSHSVISQTSFRAEYKATGGPAVQKPVKQVFDITYTEGEA 691

QY 542 QKE-----NGIYSVTFTLLSGSPRRFRKRVVETIQALLSTHDPAPAAQLSDTTN 590

DB 692 SPRDGGGGGYSVTFTLISGSPRRFKRWETIOAQLSTHDQSPVQALADEKN 746
 RESULT 7
 Q9VUV4 PRELIMINARY; PRT; 793 AA.
 ID Q9VUV4
 AC Q9VUV4
 DT 01-MAY-2000 (TrEMBLrel. 13, Created)
 DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
 DT 01-WAR-2003 (TrEMBLrel. 23, Last annotation update)
 DE CG6114 protein.
 GN Drosophila melanogaster (fruit fly).
 OS Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
 OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 OC Ephydroidea; Drosophilidae; Drosophila.
 OX NCBI_TaxID=7227;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC MEDLINE=20196006; PubMed=10731132;
 RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Cocayne J.D.,
 RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
 RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
 RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
 RA Brandon R.C., Rogers Y.-H.C., Blazek R.G., Champe M., Pfeiffer B.D.,
 RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
 RA Abil J.F., Abayani A., An H.-J., Andrews-Franknoch C., Baldwin D.,
 RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
 RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
 RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brothier P.,
 RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
 RA Cherry J.M., Cawley S., Dahle C., Davenport L.B., Davies P.,
 RA de Pablo B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
 RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
 RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
 RA Foslter C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glaeser K.,
 RA Glöck A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
 RA Harris N.L., Harvey D., Helman T.J., Hernandez J.R., Houck J.,
 RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
 RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
 RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
 RA Laško P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
 RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
 RA Merkulov G., Mileshina N.V., Mobarry C., Morris J., Moshrefi A.,
 RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
 RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,
 RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
 RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
 RA Shue B.C., Siden-Kiamos I., Simpson M., Skupeki M.P., Smith T.,
 RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
 RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
 RA Wang Z.-Y., Waesman D.A., Weinstock G.M., Weissbach J.,
 RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
 RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
 RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
 RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.
 FT "The genome sequence of *Drosophila melanogaster*."
 RL Science 287:2185-2195 (2000).
 RN [2]
 RP SEQUENCE FROM N.A.
 RA Celniker S.E., Adams M.D., Kronmiller B., Wan K.H., Holt R.A.,
 RA Evans C.A., Cocayne J.D., Amanatides P.G., Brandon R.C., Rogers Y.,
 RA Banzon J., An H., Baldwin D., Banzon J., Beeson K.Y., Busam D.A.,
 RA Carlson J.W., Center A., Champe M., Davenport L.B., Dietz S.M.,
 RA Dodson K., Dorsett V., Doup L.E., Doyle C., Dresnek D., Farfan D.,
 RA Ferreira S., Frise E., Galle R.F., Garg N.S., George R.A.,
 RA Gonzalez M., Houck J., Hoskins R.A., Hostin D., Howland T.J.,
 RA Ibegwam C., Jalali M., Kruse D., Li P., Mattei B., Moshrefi A.,
 RA McIntosh T.C., Moy M., Murphy B., Nelson C., Nelson K.A., Nunco J.,
 RA Pacleb J., Paragas V., Park S., Patel S., Pfeiffer B.,
 RA Phouanavong S., Pittman G.S., Puri V., Richards S., Scheeler F.,

RA Stapleton M., Strong R., Svirskas R., Tector C., Tyler D.,
 RA Williams S.M., Zaveri J.S., Smith H.O., Venter J.C., Rubin G.M.;
 RT "Sequencing of *Drosophila melanogaster* genome";
 RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
 RN [3]
 RP SEQUENCE FROM N.A.
 RA Misra S., Crosby M.A., Matthews B.B., Bayraktaroglu L., Campbell K.,
 RA Hradecky P., Huang Y., Kaminker J.S., Prochuk S.E., Smith C.D.,
 RA Tupy J.L., Bergman C., Berman B., Carlson J.W., Celniker S.E.,
 RA Clamp M., Drysdale R., Emmert D., Frise E., de Grey A., Harris N.,
 RA Kronmiller B., Marshall B., Millburn G., Richter J., Russo S.,
 RA Searle S.M.J., Smith E., Shu S., Smutniak F., Whitfield E.,
 RA Ashburner M., Gelbart W.M., Rubin G.M., Mungall C.J., Lewis S.E.;
 RT "Annotation of *Drosophila melanogaster* genome";
 RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
 RN [4]
 RP SEQUENCE FROM N.A.
 RA Adams M.D., Celniker S.E., Gibbs R.A., Rubin G.M., Venter C.J.;
 RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
 RN [5]
 RP SEQUENCE FROM N.A.
 RA FlyBase;
 RL Submitted (SEP-2002) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AE003529; AAF49569.2; -
 DR HSSP; Q63450; 1A06.
 DR FlyBase; FBgn0036544; CG6114.
 DR InterPro; IPR000719; Prot_kinase.
 DR InterPro; IPR002290; Ser_thr_kinase.
 DR InterPro; IPR001245; Tyr_kinase.
 DR Pfam; PF00069; pkinase; 1.
 DR ProDom; PD000001; Prot_kinase; 1.
 DR SMART; SM00220; S_TKc; 1.
 DR SMART; SM00219; TyrKc; 1.
 DR PROSITE; PS00011; PROTEIN_KINASE_DOM; 1.
 DR PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
 KW ATP-binding; Transferase.
 SQ SEQUENCE 793 AA; 87692 MW; DC0F371B489FEF48 CRC64;
 Query Match 53.7%; Score 1712; DB 5; Length 793;
 Best Local Similarity 51.0%; Pred. No. 2.1e-119; Indels 142; Gaps 19;
 Matches 370; Conservative 80; Mismatches 133;
 QY 10 LKLEHPEHLKLDVYENKYLVLVLEHVSQGGELFDYLVKKGRLTPKEARFFRQIISAL 69
 DB 1 MKLIDHPHVLGSDVYENKYLVLVLEHVSQGGELFDYLVKKGRLTPKEARFFRQIISAL 60
 QY 70 DFCSHSICHRLDKPENLLDKNRIADFGMASLQVDSLLTSCGSPHYACPEVIRG 129
 DB 61 DFCSHSICHRLDKPENLLDKNRIADFGMASLQVDSLLTSCGSPHYACPEVIRG 120
 QY 130 EKYDGRKADVWSCGVILFALLVAGALPDDNNLROLLEKVKGVFHPHFIPDCQSLRG 189
 DB 121 EKYDGRKADVWSCGVILFALLVAGALPDDNNLROLLEKVKGVFHPHFIPDCQSLRG 180
 QY 190 MIEVDAARRLTLEHIQKHIVY-IGGKNEPEQPIPRKVQIRSLPSLEIDIPDVLDSMHS 248
 DB 181 MIEVDPRLRLFLASINRHPWVTAGCKGELELPMWVQTHVPTATVDPDLVNAICS 240
 QY 249 LGCFRDRNKLQDLSEENOEKMIYFLLDRKRYPSQEDD---LPPRNE---IDPPR 302
 DB 241 LGCFCFEKELIQELLSSSHNTEKVIYFLLERKRRPALEDDDEIAQKRSBELDAVDPPR 300
 QY 303 KRVDSPLN-----RHGKRPERKSMVLSVTD 330
 DB 301 KRLDTCRINGNAPSYGQISEGSPFLTPRQAFNFRSVSTNHRORRPTVTSSVRS--SS 359
 QY 331 GGSVPARRAITEMAQ-----HGQRSSISGASGLSTSLSS--- 367
 DB 360 YHSPTRCNPSMSSAQQANAI SRPSSPAAGTRHSTGYDRDR--SGHHSVSRTFSHSSQK 417
 QY 368 -----PRVTPHPSPR-----GSP-----LPTPKGTP----- 388
 DB 418 SIEGDVVVVREPRIERDRLSLQERGGSPDRGDCGIPPGSGGNSGSGSTASPSVHRA 477

```

QY 389 -----VHTPKESPAGTNPPT-----PSSP--SVGGVPMARLNSIKNSFL 427
Db 478 NSGTIALSMFHDSDNSVNPNGSPMWNSSPGMPSFCNTPGQQLWKRLTNKNSFL 537
QY 428 GSPFRHRKRLQVPTPEMSNLTPESSPELAKXWFGNFISLEKEEQIFVWIKDKPLSSIK 487
Db 538 GSPFRHRKMQVSADE--VHLTPSSPELTKRSGWFGNLTTEKDETFTILVKGKPIATVK 595
QY 488 ADIVHAFSLISLSHVSQTSFRAEYKATG- GPAVFQKPVKFOVDIT--YTEGGEAQKE 544
Db 596 AHLHAFSLMAELSHSVSPTSFVYKRNNGPVMFORHYKFOVDISAICKQGDADM- 654
QY 545 NGIYSVTFTLLSGSPRRKRVETITQALLSTH-----DPP-----AAQHLSDTTNMEMM 595
Db 655 --LPAITFTLLSGNRRFRICEHIQSQVCKRPGPSSPTVTISVQAVSESSCGSVS 712
QY 596 TGRLS 600
Db 713 SERLS 717

RESULT 8
O61298 PRELIMINARY; PRT; 698 AA.
AC O61298;
DT 01-AUG-1998 (TremBLrel. 07, Created)
DT 01-AUG-1998 (TremBLrel. 07, Last sequence update)
DT 01-MAR-2003 (TremBLrel. 23, Last annotation update)
DE HRP0PK-1 protein.
GN HRP0PK-1.
OS Halocynthia roretzi (Sea squirt).
OC Eukaryota; Metazoa; Chordata; Urochordata; Ascidiacea;
OC Stolidobranchia; Pyruidae; Halocynthia.
OX NCBI_TaxID=7729;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Egg;
RX MEDLINE=98404280; PubMed=9767157;
RA Sasakura Y., Ogasawara M., Makabe K.W.;
RT "Maternally localized RNA encoding a serine/threonine protein kinase in the ascidian, Halocynthia roretzi."
RL Mech. Dev. 76:161-163(1998).
CC -1- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
DR EMBL; AB014885; BAA28663.1; -.
DR HSSP; O63450; 1A06.
DR InterPro; IPR000719; Prot_kinase.
DR InterPro; IPR002290; Ser_thr_kinase.
DR Pfam; PF00069; pkinase; 1.
DR PRINTS; PR00109; TYRKINASE.
DR ProDom; PD000001; Prot_kinase; 1.
DR SMART; SM00220; S_TKc; 1.
DR PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
DR PROSITE; PS00011; PROTEIN_KINASE_DOM; 1.
DR PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
DR ATP-binding; Kinase; Serine/threonine-protein kinase; Transferase.
SQ SEQUENCE 698 AA; 78607 MW; B29B1751B83200BF CRC64;

Query Match 50.1%; Score 1596.5; DB 5; Length 698;
Best Local Similarity 53.5%; Pred. No. 7.7e-111;
Matches 329; Conservative 98; Mismatches 121; Indels 67; Gaps 16;

QY 2 KVEREIALKLEIHPHLKLDHVVENKYLVLVLEHVSGLLEFDYLVKKGRLTTPKEARKF 61
Db 57 KVEREIALKLEIHPHLGLDHDVVENKYLVLVLEHVSGLLEFDYLVQKGLIIPREARRF 116
QY 62 FQRIISALDFCHSHSICHRDLKPENLLDEKNNIRIADFGMASLQVGSLLFETSCGSPHY 121
Db 117 FQRIISAVDYCHNVCHRDLPENLLDEKNNIKVADFGMASLQPEGLFLETSCGSPHY 176
QY 122 ACPEVIREKTDGRKADYVSCGVILFALLVGLALPDDNLRQLLEKVKRGVYFHPHPIPP 181

```

```

Db 177 ACPEVIRGERYDGRADVMSCGVILFALLVGLALPDDNLRQLLEKVKRGVYFHPFPVP 236
QY 182 DCQSLRGMLEVDAAERLTLHLEHIOKHIIWYIGKNE-----PEPEQPIPRKVQIRSLPS 234
Db 237 DAQNLRGMIDVRPDKRLSLQQLQHPWMPRPGSNSVEGVLVTPDPVPV-----IDCVPLPE 293
QY 235 LEDIDPDVLDMSHSLGCFRDRNKLQDLLESEBENQEKMIYFLLLDKRYPS-QEDEDLP 293
Db 294 EESVDPDVLASMTSLGCFCKEKLAKNLITEQNTKVVVYVYMLLRKKRYPSFDDADSL 353
QY 294 PRNEIDPPKRVQD--SPMLNRHG-----KRRPBRK-SMEVLSDTGGSPVPARRAEMAQ 345
Db 354 PCXHPDAPKRVDSSTSLSSNGDDWCNVP1PORKMASLCLTDSSTSLSRK--KSTE 411
QY 346 HGORSSISGASGL-----STSPLSRPTVTPHSPRGSPPLTPKGTGVH--TPKESP 396
Db 412 THORSLSGESSRLVNCNISDTKAEKSRI-----NGTPVRGTTCCSN 457
QY 397 AGTPNTPPSPSVGGVPMARLNSIKNSFLGSPFRHRKRLQVPTPEMSNLTPESSPEL 456
Db 458 QVPVQINTPASPN-----PWRQLASLKNFTMGSPFRHRKMQAPSSDEDVEN-QGNSSEL 512
QY 457 AKKSWFGNFISLE-----KEEQIFVVIKDKPLSSIKADIVHAFSLISLSHVSQTS 509
Db 513 SKRSWFGNFMSSRYSTHCHDELFPYAIAYKNRTLSNVKSELVHAFSLIPNLTHSMVSPTR 572
QY 510 FRAEYKATG-GPAVF-QKPVKFOVDITVTEGGEAQEN-----GIYSVTFTLLSGP 558
Db 573 FRCDYKSSGTSTTSVFHQRISIKFQVDLIQHSLSLDROENGKKPSSQTVGTSFTIAFSLISGP 632
QY 559 SRRFRKRVETITQAL 573
Db 633 IRRYKRVLELLQMQM 647

RESULT 9
Q19469 PRELIMINARY; PRT; 914 AA.
ID Q19469;
DT 01-NOV-1996 (TremBLrel. 01, Created)
DT 01-OCT-2001 (TremBLrel. 18, Last sequence update)
DT 01-MAR-2003 (TremBLrel. 23, Last annotation update)
DE F15A2.6 protein (Serine/threonine kinase SAD-1).
GN F15A2.6.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC Rhabditidae; Peloderinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
RP SEQUENCE FROM N.A.
RA Gregory J.;
RL Submitted (MAR-1996) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=99069613; PubMed=9851916;
RA none;
RT "Genome sequence of the nematode C.elegans: A platform for investigating biology."
RL Science 282:2012-2018(1998).
RN [3]
RP SEQUENCE FROM N.A.
RA Crump J.G., Zhen M., Jin Y., Bargmann C.;
RT "The SAD-1 kinase regulates presynaptic vesicle clustering."
RL Submitted (OCT-2000) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
DR EMBL; Z70207; CAA94127.2; -.
DR EMBL; AF316542; AAG50270.1; -.
DR HSSP; Q63450; 1A06.
DR WormPep; F15A2.6; CE28218.
DR InterPro; IPR000719; Prot_kinase.
DR InterPro; IPR002290; Ser_thr_kinase.
DR InterPro; IPR006025; Zn_Wipeptase.
DR Pfam; PF00069; pkinase; 1.

```

```
DR ProDom; P0000001; Prot_kinase; 1.
DR SMART; SMO0220; S_TK; 1.
DR PROSITE; PS00107; PROTEIN KINASE ATP; 1.
DR PROSITE; PS00111; PROTEIN_KINASE_DOM; 1.
DR PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
DR PROSITE; PS00142; ZINC_PROTEASE; 1.
KW ATP-binding; Kinase; Serine/threonine-protein kinase; Transferase.
SQ SEQUENCE 914 AA; 100840 MW; FDE311D31B249D3E CRC64;

Query Match      48.2%; Score 1536; DB 5; Length 914;
Best Local Similarity 49.0%; Pred. No. 3.7e-106;
Matches 340; Conservative 71; Mismatches 133; Indels 150; Gaps 17;

Qy 2 KVEREIAIKLIEHPVHLKLDVYENKYLVLVLEHVSQGLFDYLVKKGRLLTPKEARKF 61
Db 90 KVEREIAIMKLEIHPVHLHLVYENKYLVLVLEHVSQGLFDYLVKGRLLMSKEARKF 149
Qy 62 FROIISALDFCHSHSICHRLDKPENLLLDKKNIRIADFGVASIQVDSLETSCGSPHY 121
Db 150 FROIISALDFCHAHNICHRLDKPENLLLDERNIKVADFGVASIQVDSLETSCGSPHY 209
Qy 122 ACPEVIRGEKDGKADYWSCVILFALLVGAIPDDDNLRKLEKVKRGVFMHPTFP 181
Db 210 ACPEVIRGEKDGKADYWSCVILYALLVGAIPDDDNLRKLEKVKRGVFMHPTFP 269
Qy 182 DCQSLRLGMIEVDAAARLTLEHIQKHIYIG-KGNEPEPEQPIPKVQIRLSLELDIDP 240
Db 270 DVQSLRLAMIEVDPKRYSLADVFHPVSGTGTAKDPELELPMSQVQVTHVIPGEDSIDP 329
Qy 241 DVLDSMHSIGCFDRNKLQDLLSEENQEKMIYELLIDRKYRPSQDELDLPNNEI-- 298
Db 330 DVLRHMCIGCFDKQKQKINELLSPKHNTKRVYFLLLDKRRRPAQEDD-----TEIVL 384
Qy 299 -----DPPRKRVDSPLMLN-----HGKRR-----PE 319
Db 395 RGAQNDPPKKTDSRTSRYPMGSIADGSPINPKTYGNQKSGRHSILGGSPTESPR 444
Qy 320 RKSMEVLSVTDGGS-----PVPARRAIEMAQH----- 346
Db 445 SSTRDLFGSSSGSYARAGEDDRGRSASRSTNSYHYVTPVDPTQLAEAAHVRDAQE 504
Qy 347 -----CQSRSTSGASSGLSTPLSSPRVTPHPSRGSLPTPKTPVHTP 392
Db 505 RRESRDSGRGSRKESKDRSDKSASSCKNDASTSVPH-----KYSPPSVW 553
Qy 393 KES---PAGTNPPTPSS-----PSVGGV--PWRARLNSIKNSFLGSPRFRHRRKLIQ 438
Db 554 SESVVSSTWNSTNSSTNSLIAGNSQTSIGTSGPWRKLNINIKNSFLGTPFRHRRK-- 611
Qy 439 VPTPEEMNSLTPESSPE-----LAKKSGWGNF-----ISLEKEEQIFVVIKDKPLSSI 486
Db 612 -----MSGTAESDSQSDSDMTDLVKKSWFGSLASSMSVERDDTHCVVPQGTLSNI 665
Qy 487 KADIVHAFSLPSLSHVSQTSFRAEYK--ATGGPAVFQKPVKQVDITYTE-----GG 539
Db 666 KAEIRAFQLHLSHVSQVQNCVFVEYKRGTPVGGVSFRGKMNVDIIPSPQOQVIAG 725
Qy 540 EAQKENGYSVTFLLSGPSRRFRKRVETIOAQL 573
Db 726 ETP-----TVVQVFLVLAGVFRFKELVEHLSAIL 755

RESULT 10
Q96AV4
ID Q96AV4 PRELIMINARY; PRT; 473 AA.
AC Q96AV4;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DE 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE Hypothetical protein.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Lymph;
RA Strausberg R.;
RL Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC016681; AAH16681.1; -.
DR InterPro; IPR000449; UBA_domain.
KW Hypothetical protein.
SQ SEQUENCE 473 AA; 50750 MW; D27DAD437DA98AA9 CRC64;

Query Match      45.2%; Score 1439.5; DB 4; Length 473;
Best Local Similarity 67.5%; Pred. No. 2.5e-99;
Matches 287; Conservative 38; Mismatches 37; Indels 63; Gaps 5;

Qy 229 IRSIPLSELDIDPDLDSMHSIGCFDRNKLQDLLSEENQEKMIYELLIDRKYRPSQE 288
Db 1 MRSUPSNGEIDPDVLESMAISLGCGRDRERLHRLSEENQEKMIYELLIDRKYRPSCE 60
Qy 289 DEOLPPRNEIDPPRKRVDSPLMLNRHGKRRPERKSMELSVTD---CGSPVPARRAIEMAQ 345
Db 61 DQDLPPNDVDPPRKRVDSPLMLSRHGKRRPERKSMELSVITDAGGGSPVPTREALEMAQ 120
Qy 346 HGQSRSRISGASSGLSTPLSSPR----- 369
Db 121 HSQSRSRVSGASTGLSSPLSSPRSPVFSFSPFGAGDEARGGSPSTKTQTLPSRGPGR 180
Qy 370 -----VTPHPSRGSLPTPKG-----TPVHTPKESPAGTNPPTPSSP--SVG 411
Db 181 GGAGEQPPPSAASATPLPGPPGPRSSGGTPLHSPLHTPRASPTGTTPPPSPGGGVG 240
Qy 412 GVPWRARLNSIKNSFLGSPRFRHRRKLVPTPEEMNSLTPESSPELAKKSWFGNFIISLEKE 471
Db 241 GAWRSLNSIRNSFLGSPRFRHRRKLVPTAEEMSSITPESSPELAKKSWFGNFIISLDKE 300
Qy 472 EQIPVVIKDKPLSIKADIVHAFSLPSLSHVSQTSFRAEYKATGGPAVFQKPVKPVQV 531
Db 301 EQIPVVIKDKPLSIKADIVHAFSLPSLSHVSQTSFRAEYKATGGPAVFQKPVKPVQV 360
Qy 532 DITVTEGEAQAQE-----NGIYSVTFTLLSGPSRRFRKRVETIOAQLSTHDPAAQHL 585
Db 361 DISSEGEPEPSPRDGGGGGIYSVTFLLSGPSRRFRKRVETIOAQLSTHDPAAQHL 420
Qy 586 SDDTN 590
Db 421 ADEKN 425

RESULT 11
Q95T82
ID Q95T82 PRELIMINARY; PRT; 701 AA.
AC Q95T82;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE GH13047p.
GN CG6114.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Berkely;
RA Stapleton M.; Brokstein P.; Hong L.; Agbayani A.; Carlson J.;
RA Champe M.; Chavez C.; Dorsett V.; Farfan D.; Frise E.; George R.;
RA Gonzalez M.; Guarin H.; Li P.; Liao G.; Miranda A.; Mungall C.J.;
RA Nuncio J.; Pacle J.; Paragas V.; Park S.; Phouanavong S.; Wan K.;
RA Yu C.; Lewis S.E.; Rubin G.M.; Celniker S.;
RL Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY060288; AAL25327.1; -.
DR FlyBase; FBgn0036544; CG6114.
```

DR InterPro: IPR000719; Prot kinase.
DR Pfam: PF00069; pkinase; 1.
DR ProDom: PD000001; Prot kinase; 1.
DR PROSITE: PS50011; PROTEIN_KINASE_DOM; 1.
KW ATP-binding; Transferase.
SQ SEQUENCE 701 AA; 76973 MW; 87B30879883B2425 CRC64;

Query Match 39.1%; Score 1248; DB 5; Length 701;
Best Local Similarity 44.9%; Pred. No. 8.8e-85;
Matches 284; Conservative 76; Mismatches 131; Indels 142; Gaps 19;

QY 102 MASLQVDSLLTSCGSPHYACPEVIRGEKYDGRKADVSCGVILFALLVGVLPFDDNL 161
DB 1 MASLQVDSLLTSCGSPHYACPEVIRGEKYDGRKADVSCGVILFALLVGVLPFDDNL 60

QY 162 ROLLEKVKRGVEMPHFIPDPDCOSLLRGMLTEVDAARELTLEHIQKHLY-IGCKNPEPE 220
DB 61 ROLLEKVKRGVEMPHFIPDPDCOSLLRGMLTEVDAARELTLEHIQKHLY-IGCKNPEPE 120

QY 221 QP1PRKQVIRSLPLEDIDPDVLDMSHSLGCFDRNKLQDLSEENQEKMIYFLLDR 280
DB 121 LPMMEVQTHVIPATAVDPVNAICSLGCFEKEKLIQELSSSHNTKVIYFLLER 180

QY 281 KERYPSDEDED-LPRNE---IDPRKRVDSPLN----- 311
DB 181 KRRRPALEDDEDEIAQKSRSELDAVDPKRLDTCRNGTNPASVYQISEGSLTPRRQAF 240

QY 312 -----RHGKRRPRKSNVSLVTDGSPVPARRAIEAQ----- 345
DB 241 NFRSYSTRNHQRSPFTVTSSVRS-SYHSPTRCNSPMSSAQQAIAIRSPSPAAGTR 299

QY 346 ---HGQSRISGASSGLTSPSS-----PRVTPHPSPR-----GSP--- 380
DB 300 HSTYGDRLR-SGHSSVSRTSPSSQKSTEGDVVVVREPRIERDLSLQERGSGSPDR 357

QY 381 -LTPKGT-----VHTKPSAGTNPT----- 404
DB 358 GDCGIPGSPGSGSGTSASPSVHRANSQPTIAISMFHDPSNSVVPNGSPMNNSS 417

QY 405 ---PSSP---SVGVVPMRRLNSIKNSFLGSPRRHRLQVPTPEMSNLTPESSPELAKK 459
DB 418 PGMGSPONTGGQLWKRLTNIKNSFLGSPRRHRLQVPTPEMSNLTPESSPELTKR 475

QY 460 SWFNFISSLEKEQIFVYIKDKPLSSIKADIVHAFSLPSLSHVSISQTSFRAEYKATG- 518
DB 476 SWFGNLTTEKDEFTILVKGKPIATVKAHILHAFSLMAELSHSVSPTSFVEYKENG 535

QY 519 GPVAFQPKVQVDIT--YTGGEAQAENGYSVTFLLSGPSRRPRKRVETIIQAQLLT 576
DB 536 GPVAFQPKVQVDISAIKQGDADM---LFAITFTLLSGNIRPRRICEHIOQVCSK 592

QY 577 H-----DPP-----AAQHLSDTTNCWEMMTGRLS 600
DB 593 RFPGPSPTVTSTVQAVSESSSCGVSSERLS 625

RESULT 12
Q8SSX5 PRELIMINARY; PRT; 833 AA.
AC Q8SSX5;
DT 01-JUN-2002 (TrEMBLrel. 21, Created)
DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Putative serine/threonine protein kinase.
OS Dictyostelium discoideum (Slime mold).
OC Eukaryota; Mycetozoa; Dictyosteliida; Dictyostelium.
OX NCBI_TaxID=44689;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=AX4;
RA Gloeckner G., Eichinger L., Szafranski K., Pachebat J., Dear P.,
RA Lehmann R., Baumgart C., Parra G., April J.F., Guigo R., Kumpf K.,
RA Tunggal B., Cox E., Quail M.A., Platzer M., Rosenthal A., Noegel A.A.,

RT "Sequence and Analysis of Chromosome 2 of Dictyostelium.";
RL Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.
CC -!- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
DR EMBL: AC115685; AAL92711.1; --
DR HSSP: P24941; IBOU.
DR InterPro: IPR000719; Prot kinase.
DR InterPro: IPR002290; Ser_Thr_kinase.
DR Pfam: PF00069; pkinase; 1.
DR ProDom: PD000001; Prot kinase; 1.
DR SMART: SM00220; S_TKc_1.
DR PROSITE: PS50011; PROTEIN_KINASE_DOM; 1.
DR PROSITE: PS00108; PROTEIN_KINASE_ST; 1.
KW ATP-binding; Kinase; Serine/threonine-protein kinase; Transferase.
SQ SEQUENCE 833 AA; 93467 MW; 51B64036C90CF94 CRC64;

Query Match 24.1%; Score 769.5; DB 5; Length 833;
Best Local Similarity 29.5%; Pred. No. 6.6e-49;
Matches 190; Conservative 88; Mismatches 154; Indels 211; Gaps 15;

QY 2 KVEREIALKLIEHPHVLKLDHVENKKYLVLEHVSQGLFDYLVKGRITPKEARXP 61
DB 54 KIEREIVLMKLDHNPAMQVEVETSKYLFILVEVEGGLFDYLVKGLSEGEALFF 113

QY 62 FQOITSALDFCHSHSICHRDLKPEMLLDEKNIRIADFGMASLQVDSLLTSCGSPHY 121
DB 114 FQOIIIGLEYCHSRNICHRDLKPEMLLSDGDKRIKIADFGMSIVRKOMLLHTSCGSPHY 173

QY 122 ACEPVIRGEKYDGRKADVSCGVILFALLVGVLPFDDNLRLLEKVGKGVFHPHFIPP 181
DB 174 ASPEVVGSDYDQKADVSCGVILYALTUTKLPFDDENIRLLNKKVKGATSMPPYIHK 233

QY 182 DQOSLRGMIEVDAAARLTLEHIQKHLYIGGKNEPEPEQIPRKVQIRSLPLEDIDPD 241
DB 234 DAQDLITKMLTVDPKSIKEIKHPWFVSNFQKATPV-EEINASPLVDYSQIDED 292

QY 242 VLDMSHSLGCFDRNKLQDLSEENQEKMIYFLLDRKRY----- 284
DB 293 IFRSLMALG-VGTIDEVKQQLVSNQKS-ATLIYRLLERKKKFDSDVNKYGYKPKETRRN 350

QY 285 ----- 284
DB 351 SLSDMSLKKIFSGSNN 410

QY 285 -----PSQEDDLP 293
DB 411 NNN 470

QY 294 PRNEIDPPKRVDSPLNHRGKERP-----ERKSMEVLSVTDGSGSPV 336
DB 471 PSNSAN--NMAIQOPTINNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 525

QY 337 ARRAIEMAOHQGRSRSISG-----ASSGLSTPSLSPRVTPHPSPRGSPPLP 382
DB 526 LKQALQO-HHQOQQOQFNGNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 584

QY 383 TPKGTVHTPKESPAG--TPNPTPPSPSVGGVPMRRLNSIKNSFLGSPRRHRLQV 440
DB 585 SPQLSAIKPDHYQRRGSMTASTNPATSPMS-----HRGKTSSP 623

QY 441 -----TPEMSNTLPE-----SPELAKKSWFGNFIS 467
DB 624 IEITSKVRKRLKISEQSN-TPNSPIIGSSP---KXSWFSYFFS 662

RESULT 13
Q8TB60 PRELIMINARY; PRT; 170 AA.
ID Q8TB60
AC Q8TB60;
DT 01-JUN-2002 (TrEMBLrel. 21, Created)
DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE Similar to serine/threonine kinase 29.
OS Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

NCBI_TaxID=9606;

[1] SEQUENCE FROM N.A.

RP TISSUE=Eye;

RA Strausberg R.;

RL Submitted (FEB-2002) to the EMBL/GenBank/DBJ databases.

DR EMBL; BC024291; AAH24291.1; --

KW Kinase.

SQ SEQUENCE 170 AA; 18488 MW; 55D569F31P5E651A CRC64;

Query Match 22.5%; Score 718; DB 4; Length 170;

Best Local Similarity 99.3%; Pred. No. 5.6e-46;

Matches 142; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 445 MSNLTPESSPELAKSWFGNFTLSKEQIFVVIKDKPLSSIKADIVHAFISIPLSHSV 504

DB 1 MSNLTPESSPELAKSWFGNFTLSKEQIFVVIKDKPLSSIKADIVHAFISIPLSHSV 60

QY 505 ISQTSFRAEYKATGGPAVFKQVQVDITYTEGGEAKENGIVSVTTLTSGPSRRPKR 564

DB 61 ISQTSFRAEYKATGGPAVFKQVQVDITYTEGGEAKENGIVSVTTLTSGPSRRPKR 120

QY 565 VVETIQALLSTHDPAPQAHLSD 587

DB 121 VVETIQALLSTHDPAPQAHLSE 143

RESULT 14

Q98SC8

ID Q98SC8 PRELIMINARY; PRT; 401 AA.

AC Q98SC8;

DT 01-OCT-2001 (TrEMBLrel. 18, Created)

DT 01-OCT-2001 (TrEMBLrel. 18, Last sequence update)

DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)

DE SNF1-related protein kinase.

GN KIN(SNF2).

OS Guillardia theta (Cryptomonas phi).

OC Eukaryota;

OC Nucleomorph.

OC Guillardia theta (Cryptomonas phi).

OC Eukaryota;

OC Nucleomorph.

OC Guillardia theta (Cryptomonas phi).

OC Eukaryota;

OC Nucleomorph.

OC Guillardia theta (Cryptomonas phi).

OC Eukaryota;

OC Nucleomorph.

OC Guillardia theta (Cryptomonas phi).

OC Eukaryota;

OC Nucleomorph.

OC Guillardia theta (Cryptomonas phi).

OC Eukaryota;

OC Nucleomorph.

OC Guillardia theta (Cryptomonas phi).

OC Eukaryota;

OC Nucleomorph.

OC Guillardia theta (Cryptomonas phi).

OC Eukaryota;

OC Nucleomorph.

OC Guillardia theta (Cryptomonas phi).

OC Eukaryota;

OC Nucleomorph.

OC Guillardia theta (Cryptomonas phi).

OC Eukaryota;

OC Nucleomorph.

OC Guillardia theta (Cryptomonas phi).

OC Eukaryota;

OC Nucleomorph.

OC Guillardia theta (Cryptomonas phi).

OC Eukaryota;

OC Nucleomorph.

OC Guillardia theta (Cryptomonas phi).

OC Eukaryota;

OC Nucleomorph.

OC Guillardia theta (Cryptomonas phi).

OC Eukaryota;

OC Nucleomorph.

OC Guillardia theta (Cryptomonas phi).

OC Eukaryota;

OC Nucleomorph.

OC Guillardia theta (Cryptomonas phi).

Db 119 FFHEIISGLEVCHKHCHRICRDLKLENILLDKLQKIADFGWASLSIPNIMLTKFCGSPH 178

QY 121 YACPEVIRGEKYDGRKADVSCGVILFALLVGLPDD--DNLRLLEKVKRGVFMHPHF 178

Db 179 YASPEVVSNEFYNGIKADIWSCGIIYALVVGKLPYDEENDNMKLFNKIRFEPPIPRI 238

QY 179 IPPQCQILLRGMIEVDAARLLTLEHIQKIWIYIGCKNEPEPEQIPRKVQIRSLP----- 233

Db 239 IHPDRLIRALLTINPQKITIDIKKHPY---KSPLPES---CRISIQNISFASVQ 292

QY 234 -SLEDIDPDVLDMSHSLGCFRDRNKLQDLSEENQKMYIFLLDRKERYPSQEDDL 292

Db 293 NPISNPDEIISFLLQLVKDKKTLGCLSTKPSFLVLYRQLEWKRKMKMDSVRSNLI 352

QY 293 PPRNEIDPPKRVDSPLNRHCK 315

Db 353 ---NEANFKRK-----LFQOGK 367

RESULT 15

Q8RWD2

ID Q8RWD2 PRELIMINARY; PRT; 535 AA.

AC Q8RWD2;

DT 01-JUN-2002 (TrEMBLrel. 21, Created)

DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)

DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)

DE Putative SNF1-related protein kinase.

GN AT3G01090.

OS Arabidopsis thaliana (Mouse-ear cress).

OC Eukaryota;

OC Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;

OC eurosids II; Brassicales; Brassicaceae; Arabidopsi.

OC NCBI_TaxID=3702;

OX [1]

RN SEQUENCE FROM N.A.

RP Southwick A., Karlin-Neumann G., Nguyen M., Lam B., Miranda M.,

RA Palm C.J., Bowser L., Jones T., Banh J., Carninci P., Chen H.,

RA Cheuk R., Chung M.K., Hayashizaki Y., Ishida J., Kamiya A., Kawai J.,

RA Kim C., Lin J., Liu S.X., Narusaka M., Pham P.K., Sakano H.,

RA Sakurai T., Satou M., Seki M., Shinn P., Yamada K., Shinozaki K.,

RA Ecker J., Theologis A., Davis R.W.,

RA Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.

CC -1- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.

DR EMBL; AY093170; AA013169.1; --

DR HSSP; P24941; 1BUH

DR InterPro; IPR001772; Kinase Cterm.

DR InterPro; IPR000719; Prot_kinase.

DR InterPro; IPR002290; Ser_thr_kinase.

DR InterPro; IPR000449; UBA_domain.

DR Pfam; PF02149; KAL; 1.

DR Pfam; PF00069; Kinase; 1.

DR Pfam; PF00627; UBA; 1.

DR ProDom; PD000001; Prot_kinase; 1.

DR SMART; SM00220; S_TKc; 1.

DR SMART; SM00165; UBA; 1.

DR PROSITE; PS00107; PROTEIN KINASE ATP; 1.

DR PROSITE; PS00011; PROTEIN KINASE_DOM; 1.

DR PROSITE; PS00108; PROTEIN KINASE_ST; 1.

KW ATP-binding; Kinase; Serine/threonine-protein kinase; Transferase.

SQ SEQUENCE 535 AA; 61181 MW; FFFC383223FD8317 CRC64;

Query Match 21.9%; Score 698; DB 10; Length 535;

Best Local Similarity 42.7%; Pred. No. 8.1e-44;

Matches 137; Conservative 70; Mismatches 90; Indels 24; Gaps 5;

QY 2 KVEREIAILKLEIHPHYVLKLDVYENKKYLYLVLEHVSGGELFDYLVKGRLTPEARKF 61

Db 86 KVRREIKILRLEWHPHPIRLYVETPTDYLWYVNSGELFDYIVEKRLQDEARNF 145

QY 62 FQRIISALDFCHSHSICHRLDKPENLLLDKNNIRIADFGWASLQVGDLSLETSCGSPHY 121

Db 146 FQRIISGVEYCHRMVYVHRLDKPENLLLDKNNIRIADFGWASLQVGDLSLETSCGSPHY 205


```
QY 122 ACEVIRGEKYDGRKADVNSCGVILFALLVGCALPFDDNLRQLLEKXKGVFHMHPHFIP 181
Db 206 APEVISGKLIYAGPEVDVWSCGVILYALLCGTLPFDDENIPNLFKIKIGIYTLPSHLSP 265
QY 182 DCOSLLRGMIEVDAARLTLLEHIQKHIWIYIGKNEPEPEQPIPRKVQI---RSLPSLEDI 238
Db 266 GARDLIPRLVVDPMKRVTTIPEIRQHPWF-----QAHLPRYLAVPPPTVQQAKKI 316
QY 239 DDPVLDMSHSLGCFRDRNKLQDLLSEENQEKMIYFLLLDKRYPS-----QEDDED 291
Db 317 DEBILQEVINMGF--DRNHILIESLRNRQTNDGTVTYYLILDNRFRASSGYLGAEFQETME 374
QY 292 LPPRNEIDPPEKRVDSPLNR 312
Db 375 GTPRMH---PAESVAPVSHR 392
```

Search completed: November 26, 2003, 12:01:54
Job time : 34.4013 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 26, 2003, 12:01:57 ; Search time 25.7304 Seconds
(without alignments)
4358.326 Million cell updates/sec

Title: US-10-054-579-4

Perfect score: 3188

Sequence: 1 MKVRETAILKLIIEHPHVLK.....TNCMMTGRSLKCGIIPKS 608

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 673684 seqs, 184443283 residues

Total number of hits satisfying chosen parameters: 673684

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA:

- 1: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep.*
- 2: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pep.*
- 3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep.*
- 4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep.*
- 5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep.*
- 6: /cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB.pep.*
- 7: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep.*
- 8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep.*
- 9: /cgn2_6/ptodata/2/pubpaa/US09A_PUBCOMB.pep.*
- 10: /cgn2_6/ptodata/2/pubpaa/US09B_PUBCOMB.pep.*
- 11: /cgn2_6/ptodata/2/pubpaa/US09C_PUBCOMB.pep.*
- 12: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pep.*
- 13: /cgn2_6/ptodata/2/pubpaa/US10A_PUBCOMB.pep.*
- 14: /cgn2_6/ptodata/2/pubpaa/US10B_PUBCOMB.pep.*
- 15: /cgn2_6/ptodata/2/pubpaa/US10C_PUBCOMB.pep.*
- 16: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pep.*
- 17: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep.*
- 18: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	3188	100.0	608	14	US-10-054-579-4
2	3188	100.0	664	12	US-10-288-798-18
3	3188	100.0	668	14	US-10-054-579-2
4	3188	100.0	668	15	US-10-195-072-2
5	3188	100.0	668	15	US-10-195-071-2
6	3071	96.3	674	15	US-10-283-247-2
7	3065	96.1	614	15	US-10-283-247-9
8	3065	96.1	614	15	US-10-283-247-10
9	3065	96.1	674	10	US-09-842-582-2
10	3065	96.1	674	15	US-10-283-247-7
11	3065	96.1	674	15	US-10-283-247-8
12	3050	95.7	636	15	US-10-283-247-5
13	3028	95.0	585	15	US-10-195-072-4
14	3028	95.0	585	15	US-10-195-071-4
15	3015	94.6	603	15	US-10-195-072-17

16	3015	94.6	603	15	US-10-195-071-17	Sequence 17, Appl
17	2545.5	79.8	703	12	US-10-116-326-6	Sequence 6, Appl
18	2545.5	79.8	778	12	US-10-354-358-92	Sequence 92, Appl
19	2545.5	79.8	778	12	US-10-116-326-2	Sequence 2, Appl
20	2386.5	74.9	762	12	US-10-116-326-4	Sequence 4, Appl
21	722	22.6	1518	10	US-09-801-368-152	Sequence 152, App
22	688.5	21.6	745	12	US-10-161-565-24	Sequence 24, Appl
23	688.5	21.6	745	15	US-10-195-101-36	Sequence 36, Appl
24	686.5	21.5	691	10	US-09-919-585-6	Sequence 6, Appl
25	686.5	21.5	724	10	US-09-919-585-9	Sequence 9, Appl
26	682.5	21.4	722	8	US-08-817-8328-32	Sequence 32, Appl
27	680	21.3	149	12	US-10-102-548-2	Sequence 2, Appl
28	680	21.3	149	15	US-10-102-558-2	Sequence 2, Appl
29	680	21.3	149	15	US-10-195-072-18	Sequence 18, Appl
30	680	21.3	149	15	US-10-195-071-18	Sequence 18, Appl
31	675.5	21.2	752	10	US-09-835-081-2	Sequence 2, Appl
32	671.5	21.1	688	12	US-10-161-565-28	Sequence 28, Appl
33	671.5	21.1	688	12	US-10-161-565-29	Sequence 29, Appl
34	671.5	21.1	744	10	US-09-835-081-4	Sequence 4, Appl
35	670.5	21.0	795	10	US-09-919-585-12	Sequence 12, Appl
36	670.5	21.0	795	12	US-10-161-565-25	Sequence 25, Appl
37	670.5	21.0	795	15	US-10-142-356-9	Sequence 9, Appl
38	667	20.9	1369	12	US-10-288-798-7	Sequence 7, Appl
39	663.5	20.8	729	12	US-10-161-565-26	Sequence 26, Appl
40	663.5	20.8	729	15	US-10-142-356-11	Sequence 11, Appl
41	663.5	20.8	729	15	US-10-195-101-33	Sequence 33, Appl
42	662.5	20.8	783	9	US-09-815-915-2	Sequence 2, Appl
43	662.5	20.8	783	11	US-09-823-187-90	Sequence 90, Appl
44	662.5	20.8	783	12	US-10-393-316-2	Sequence 2, Appl
45	657.5	20.6	744	10	US-09-919-585-3	Sequence 3, Appl

ALIGNMENTS

RESULT 1

US-10-054-579-4
; Sequence 4, Application US/10054579
; Publication No. US20020137913A1
GENERAL INFORMATION:
; APPLICANT: Turner, C. Alexander Jr.
; APPLICANT: Mathur, Brian
; TITLE OF INVENTION: No. US20020137913A1el Human Kinases and Polynucleotides Encoding t
; FILE REFERENCE: LEX-0300-USA
; CURRENT APPLICATION NUMBER: US/10/054,579
; CURRENT FILING DATE: 2002-01-22
; PRIOR APPLICATION NUMBER: US 60/263,378
; PRIOR FILING DATE: 2001-01-23
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 608
; TYPE: PRT
; ORGANISM: homo sapiens
US-10-054-579-4

Query Match	100.0%;	Score 3188;	DB 14;	Length 608;
Best Local Similarity	100.0%;	Pred. No. 2.5e-196;		
Matches 608;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
Qy	1	MKVRETAILKLIIEHPHVLKLDVYENKYLVLVLEHVS	GGELFDYLVKKGRLLTPKEARK	60
Db	1	MKVRETAILKLIIEHPHVLKLDVYENKYLVLVLEHVS	GGELFDYLVKKGRLLTPKEARK	60
Qy	61	FFROIISALDPCHSHSICHRLDKPENLLDDEKNIRIADFG	MASLQVGDLSLETSCGSPH	120
Db	61	FFROIISALDPCHSHSICHRLDKPENLLDDEKNIRIADFG	MASLQVGDLSLETSCGSPH	120
Qy	121	YACPEVIRGEKYDGRKADVWMSGCVILFALLVGLALP	FFDDNLRQLLEKVKRGVFFHMFIP	180
Db	121	YACPEVIRGEKYDGRKADVWMSGCVILFALLVGLALP	FFDDNLRQLLEKVKRGVFFHMFIP	180
Qy	181	PDQSLLRGMIEVDARRRLTLEHIQKHWITGGKNEPEPE	QPIPRKVOIRSLPSLEDIDP	240

181 PDCQSLRGMEVDAARLTLEHIQKHWYIGGKNEPEPEPIPRKQVIRSLPSLEIDP 240
241 DVLDMSHSLGCFRDRNKLQDLSEENQKMIYFLLIDRKERYPSQEDDLPRNEIDP 300
241 DVLDMSHSLGCFRDRNKLQDLSEENQKMIYFLLIDRKERYPSQEDDLPRNEIDP 300
301 PRKRVDSPLNLRHGRKRPKRKMEVLSVTDGSGSPVARRAIEAQAHQQRSSISGASSGL 360
301 PRKRVDSPLNLRHGRKRPKRKMEVLSVTDGSGSPVARRAIEAQAHQQRSSISGASSGL 360
361 STSPLSSPRVTPHPSPRGSPPLTPKGTVPHTPKESPAGTNPPTPPSPSVGGVPMRRLN 420
361 STSPLSSPRVTPHPSPRGSPPLTPKGTVPHTPKESPAGTNPPTPPSPSVGGVPMRRLN 420
421 SIKNSFLGSPFRHRRKLQVPTPEMSNLTPSSPDLAKKSWFGNFIKLEKEQIFVVIKD 480
421 SIKNSFLGSPFRHRRKLQVPTPEMSNLTPSSPDLAKKSWFGNFIKLEKEQIFVVIKD 480
481 KPLSSIKADIVHAFSLSPSLSHSVISQTSFRAEYKATGGPAVFQKPVKFOVDITYTEGGE 540
481 KPLSSIKADIVHAFSLSPSLSHSVISQTSFRAEYKATGGPAVFQKPVKFOVDITYTEGGE 540
541 AQKNGIYSVTFTLLSGPSRRFRKRVVETIQALLSTHDPAAQHLSDTTNCMEMMTGRLS 600
541 AQKNGIYSVTFTLLSGPSRRFRKRVVETIQALLSTHDPAAQHLSDTTNCMEMMTGRLS 600
601 KCGIIPKS 608
601 KCGIIPKS 608

RESULT 2

US-10-288-798-18

Sequence 18, Application US/10288798

Publication No. US20030207299A1

GENERAL INFORMATION:

APPLICANT: BANDMAN, Olga; NGUYEN, Dannie B;
APPLICANT: WALIA, Naxinder K.; HAPALIA, April J.A.;
APPLICANT: YAO, Monique G.; GANDHI, Ameena R.;
APPLICANT: GURURAJAN, Rajagopal.; DING, Li;
APPLICANT: PATTERSON, Chandra.; YUE, Henry;
APPLICANT: BAUGHN, Marian R.; TRIBOULEY, Catherine M.;
APPLICANT: LU, Yan.; ISON, Craig H.;
APPLICANT: AU-YOUNG, Janice.; TANG, Y. Tom;
APPLICANT: AZIMZAI, Yalda.; BURRILL, John D.;
APPLICANT: MARCUS, Gregory A.; ZINGLER, Kurt A.;
APPLICANT: LU, Dying Aina M.; LAL, Preeti G.;
APPLICANT: RAMKUMAR, Jayalaxmi.; WARREN, Bridget A.;
APPLICANT: KEARNEY, Liam.; POLICKY, Jennifer L.;
APPLICANT: THANGAVELU, Kavitha.; BURFORD, Neil

TITLE OF INVENTION: HUMAN KINASES

FILE REFERENCE: PI-0209 USA

CURRENT APPLICATION NUMBER: US/10/288,798

CURRENT FILING DATE: 2002-11-01

PRIOR APPLICATION NUMBER: PCT/US01/27219

PRIOR FILING DATE: 2001-08-31

PRIOR APPLICATION NUMBER: US 60/240,542

PRIOR FILING DATE: 2000-10-13

PRIOR APPLICATION NUMBER: US 60/238,389

PRIOR FILING DATE: 2000-10-06

PRIOR APPLICATION NUMBER: US 60/236,499

PRIOR FILING DATE: 2000-09-29

PRIOR APPLICATION NUMBER: US 60/234,902

PRIOR FILING DATE: 2000-09-22

PRIOR APPLICATION NUMBER: US 60/232,654

PRIOR FILING DATE: 2000-09-14

PRIOR APPLICATION NUMBER: US 60/231,357

PRIOR FILING DATE: 2000-09-08

PRIOR APPLICATION NUMBER: US 60/229,873

PRIOR FILING DATE: 2000-08-31

NUMBER OF SEQ ID NOS: 48

SOFTWARE: PERL Program

SEQ ID NO 18

LENGTH: 664

TYPE: PRT

ORGANISM: Homo sapiens

FEATURE:

NAME/KEY: misc.feature

OTHER INFORMATION: Incyte ID No. US20030207299A1 4022651CD1

US-10-288-798-18

Query Match 100.0%; Score 3188; DB 12; Length 664;

Best Local Similarity 100.0%; Pred. No. 2.8e-196; Indels 0; Gaps 0;

Matches 608; Conservative 0; Mismatches 0;

1 MKVEREIALKLIIEPHVLKLDVVENKYLVLVLEHVS GGELFDYLVKKGRLTPKEARK 60

57 MKVEREIALKLIIEPHVLKLDVVENKYLVLVLEHVS GGELFDYLVKKGRLTPKEARK 116

61 FFRQIISALDFCHSHSICHRLDKPENLLDEKKNIRIADFGMASLQVGDLSLETSCGSPH 120

117 FFRQIISALDFCHSHSICHRLDKPENLLDEKKNIRIADFGMASLQVGDLSLETSCGSPH 176

121 YACPEVIRGEKYDGRKADVWSCGVILFALLVGLPDDDNLRQLLEKVKRGVFMHPHFI 180

177 YACPEVIRGEKYDGRKADVWSCGVILFALLVGLPDDDNLRQLLEKVKRGVFMHPHFI 236

181 PDCQSLRGMEVDAARLTLEHIQKHWYIGGKNEPEPEPIPRKQVIRSLPSLEIDP 240

237 PDCQSLRGMEVDAARLTLEHIQKHWYIGGKNEPEPEPIPRKQVIRSLPSLEIDP 296

241 DVLDMSHSLGCFRDRNKLQDLSEENQKMIYFLLIDRKERYPSQEDDLPRNEIDP 300

297 DVLDMSHSLGCFRDRNKLQDLSEENQKMIYFLLIDRKERYPSQEDDLPRNEIDP 356

301 PRKRVDSPLNLRHGRKRPKRKMEVLSVTDGSGSPVARRAIEAQAHQQRSSISGASSGL 360

357 PRKRVDSPLNLRHGRKRPKRKMEVLSVTDGSGSPVARRAIEAQAHQQRSSISGASSGL 416

361 STSPLSSPRVTPHPSPRGSPPLTPKGTVPHTPKESPAGTNPPTPPSPSVGGVPMRRLN 420

417 STSPLSSPRVTPHPSPRGSPPLTPKGTVPHTPKESPAGTNPPTPPSPSVGGVPMRRLN 476

421 SIKNSFLGSPFRHRRKLQVPTPEMSNLTPSSPDLAKKSWFGNFIKLEKEQIFVVIKD 480

477 SIKNSFLGSPFRHRRKLQVPTPEMSNLTPSSPDLAKKSWFGNFIKLEKEQIFVVIKD 536

481 KPLSSIKADIVHAFSLSPSLSHSVISQTSFRAEYKATGGPAVFQKPVKFOVDITYTEGGE 540

537 KPLSSIKADIVHAFSLSPSLSHSVISQTSFRAEYKATGGPAVFQKPVKFOVDITYTEGGE 596

541 AQKNGIYSVTFTLLSGPSRRFRKRVVETIQALLSTHDPAAQHLSDTTNCMEMMTGRLS 600

597 AQKNGIYSVTFTLLSGPSRRFRKRVVETIQALLSTHDPAAQHLSDTTNCMEMMTGRLS 656

601 KCGIIPKS 608

657 KCGIIPKS 664

RESULT 3

US-10-054-579-2

Sequence 2, Application US/10054579

Publication No. US20020137913A1

GENERAL INFORMATION:

APPLICANT: Turner, C. Alexander Jr.

APPLICANT: Mathur, Brian

TITLE OF INVENTION: No. US20020137913A1 Human Kinases and Polynucleotides Encoding t

FILE REFERENCE: LEX-0300-USA

CURRENT APPLICATION NUMBER: US/10/054,579

CURRENT FILING DATE: 2002-01-22

PRIOR APPLICATION NUMBER: US 60/263,378

PRIOR FILING DATE: 2001-01-23

NUMBER OF SEQ ID NOS: 4

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 2

; LENGTH: 668

; TYPE: PRT

; ORGANISM: homo sapiens

US-10-054-579-2

Query Match 100.0%; Score 3188; DB 14; Length 668;

Best Local Similarity 100.0%; Pred. No. 2.8e-196;

Matches 608; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
QY 1 MKVEREAILKLIIEHPHVLKLDHVVYENKYLVLVLEHVSQGLFDYLVKKGRLTPKEARK 60
   |||||
Db 61 MKVEREAILKLIIEHPHVLKLDHVVYENKYLVLVLEHVSQGLFDYLVKKGRLTPKEARK 120

QY 61 FFRQIISALDFCHSHSICHRLDKPENLLDEKNIRIADFGWASIQVGSLLTSCGSPH 120
   |||||
Db 121 FFRQIISALDFCHSHSICHRLDKPENLLDEKNIRIADFGWASIQVGSLLTSCGSPH 180

QY 121 YACPEVIRGEKYDGRKADWMSGCVILFALLVGLPFDNDNLRLQLLEKVKRGVFMHPHIP 180
   |||||
Db 181 YACPEVIRGEKYDGRKADWMSGCVILFALLVGLPFDNDNLRLQLLEKVKRGVFMHPHIP 240

QY 181 PDCQSLLRGMIEVDAAARLTLEHIQKHIIWYIGKNEPEPEQIPRKQVIRSLPSLEDIDP 240
   |||||
Db 241 PDCQSLLRGMIEVDAAARLTLEHIQKHIIWYIGKNEPEPEQIPRKQVIRSLPSLEDIDP 300

QY 241 DVLDMSHSLGCFRDRNKLQDLLSEENQEKMIYFLLLDKERYPSQDEDLPPRNEIDP 300
   |||||
Db 301 DVLDMSHSLGCFRDRNKLQDLLSEENQEKMIYFLLLDKERYPSQDEDLPPRNEIDP 360

QY 301 PRKRVDSFPLNHRHGKRRPERKSMELSVTDGGSVPARRAIEMAHQGRSISGASSGL 360
   |||||
Db 361 PRKRVDSFPLNHRHGKRRPERKSMELSVTDGGSVPARRAIEMAHQGRSISGASSGL 420

QY 361 STSPLSSPRVTPHPSPRGSPLTPKGTVPHTPKESPAGTNPPTSSPSVGVPMRRLN 420
   |||||
Db 421 STSPLSSPRVTPHPSPRGSPLTPKGTVPHTPKESPAGTNPPTSSPSVGVPMRRLN 480

QY 421 SIKNSFLGSPRFRHRKLOVPTPEEMSNLTPESSPELAKKSWFGNFI SLEKEQIFVVIKD 480
   |||||
Db 481 SIKNSFLGSPRFRHRKLOVPTPEEMSNLTPESSPELAKKSWFGNFI SLEKEQIFVVIKD 540

QY 481 KPLSSIKADIVHAFSLPSLSHVSISQTSFRAEYKATGPAVFQKPVKFQVDITVTEGGE 540
   |||||
Db 541 KPLSSIKADIVHAFSLPSLSHVSISQTSFRAEYKATGPAVFQKPVKFQVDITVTEGGE 600

QY 541 AQKENGIVSVTFTLLSGSPRRKRVVETIQALLSTHDPAAQHLSDTTNCMMMTGRLS 600
   |||||
Db 601 AQKENGIVSVTFTLLSGSPRRKRVVETIQALLSTHDPAAQHLSDTTNCMMMTGRLS 660

QY 601 KCGIIPKS 608
   |||||
Db 661 KCGIIPKS 668
```

RESULT 4

US-10-195-072-2

; Sequence 2, Application US/10195072

; Publication No. US20030092036A1

; GENERAL INFORMATION:

; APPLICANT: Origene Technologies

; TITLE OF INVENTION: Full-length Serine Protein Kinase in Brain and Pancreas

; FILE REFERENCE: 16U 101 C2

; CURRENT APPLICATION NUMBER: US/10/195,072

; CURRENT FILING DATE: 2002-07-15

; PRIOR FILING DATE: 2001-08-16

; NUMBER OF SEQ ID NOS: 18

; SOFTWARE: PatentIn version 3.1

; SEQ ID NO 2

; LENGTH: 668

; TYPE: PRT

; ORGANISM: Homo sapiens

US-10-195-072-2

Query Match 100.0%; Score 3188; DB 15; Length 668;

Best Local Similarity 100.0%; Pred. No. 2.8e-196;

Matches 608; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
QY 1 MKVEREAILKLIIEHPHVLKLDHVVYENKYLVLVLEHVSQGLFDYLVKKGRLTPKEARK 60
   |||||
Db 61 MKVEREAILKLIIEHPHVLKLDHVVYENKYLVLVLEHVSQGLFDYLVKKGRLTPKEARK 120

QY 61 FFRQIISALDFCHSHSICHRLDKPENLLDEKNIRIADFGWASIQVGSLLTSCGSPH 120
   |||||
Db 121 FFRQIISALDFCHSHSICHRLDKPENLLDEKNIRIADFGWASIQVGSLLTSCGSPH 180

QY 121 YACPEVIRGEKYDGRKADWMSGCVILFALLVGLPFDNDNLRLQLLEKVKRGVFMHPHIP 180
   |||||
Db 181 YACPEVIRGEKYDGRKADWMSGCVILFALLVGLPFDNDNLRLQLLEKVKRGVFMHPHIP 240

QY 181 PDCQSLLRGMIEVDAAARLTLEHIQKHIIWYIGKNEPEPEQIPRKQVIRSLPSLEDIDP 240
   |||||
Db 241 PDCQSLLRGMIEVDAAARLTLEHIQKHIIWYIGKNEPEPEQIPRKQVIRSLPSLEDIDP 300

QY 241 DVLDMSHSLGCFRDRNKLQDLLSEENQEKMIYFLLLDKERYPSQDEDLPPRNEIDP 300
   |||||
Db 301 DVLDMSHSLGCFRDRNKLQDLLSEENQEKMIYFLLLDKERYPSQDEDLPPRNEIDP 360

QY 301 PRKRVDSFPLNHRHGKRRPERKSMELSVTDGGSVPARRAIEMAHQGRSISGASSGL 360
   |||||
Db 361 PRKRVDSFPLNHRHGKRRPERKSMELSVTDGGSVPARRAIEMAHQGRSISGASSGL 420

QY 361 STSPLSSPRVTPHPSPRGSPLTPKGTVPHTPKESPAGTNPPTSSPSVGVPMRRLN 420
   |||||
Db 421 STSPLSSPRVTPHPSPRGSPLTPKGTVPHTPKESPAGTNPPTSSPSVGVPMRRLN 480

QY 421 SIKNSFLGSPRFRHRKLOVPTPEEMSNLTPESSPELAKKSWFGNFI SLEKEQIFVVIKD 480
   |||||
Db 481 SIKNSFLGSPRFRHRKLOVPTPEEMSNLTPESSPELAKKSWFGNFI SLEKEQIFVVIKD 540

QY 481 KPLSSIKADIVHAFSLPSLSHVSISQTSFRAEYKATGPAVFQKPVKFQVDITVTEGGE 540
   |||||
Db 541 KPLSSIKADIVHAFSLPSLSHVSISQTSFRAEYKATGPAVFQKPVKFQVDITVTEGGE 600

QY 541 AQKENGIVSVTFTLLSGSPRRKRVVETIQALLSTHDPAAQHLSDTTNCMMMTGRLS 600
   |||||
Db 601 AQKENGIVSVTFTLLSGSPRRKRVVETIQALLSTHDPAAQHLSDTTNCMMMTGRLS 660

QY 601 KCGIIPKS 608
   |||||
Db 661 KCGIIPKS 668
```

RESULT 5

US-10-195-071-2

; Sequence 2, Application US/10195071

; Publication No. US20030096271A1

; GENERAL INFORMATION:

; APPLICANT: Origene Technologies

; TITLE OF INVENTION: Full-length Serine Protein Kinase in Brain and Pancreas

; FILE REFERENCE: 16U 101 C1

; CURRENT APPLICATION NUMBER: US/10/195,071

; CURRENT FILING DATE: 2002-07-15

; PRIOR APPLICATION NUMBER: US 09/930,181

; PRIOR FILING DATE: 2001-08-16

; NUMBER OF SEQ ID NOS: 18

; SOFTWARE: PatentIn version 3.1

; SEQ ID NO 2

; LENGTH: 668

; TYPE: PRT

; ORGANISM: Homo sapiens

US-10-195-071-2

Query Match 100.0%; Score 3188; DB 15; Length 668;

Best Local Similarity 100.0%; Pred. No. 2.8e-196;	
Matches 608; Conservative 0; Mismatches 0; Indels 0; Gaps 0;	
QY	1 MKVEREIAILKLTIEHPHVLKLDHYVYENKKYLVLVLEHVS GGELFDYLVKKGRLLTPKEARK 60
DB	
61 MKVEREIAILKLTIEHPHVLKLDHYVYENKKYLVLVLEHVS GGELFDYLVKKGRLLTPKEARK 120	
61 FFRQIIISALDFCHSHSICHRLDPENLLIDKKNIRIADFGMASLOVGS DLSLETSCGSPH 120	
DB	
121 FFRQIIISALDFCHSHSICHRLDPENLLIDKKNIRIADFGMASLOVGS DLSLETSCGSPH 180	
QY	121 YACPEVIRGEKYDGRKADVMSCGVILFALLV GALPDDDDNLRLLKLVKRGVFMHPHFIP 180
DB	
181 YACPEVIRGEKYDGRKADVMSCGVILFALLV GALPDDDDNLRLLKLVKRGVFMHPHFIP 240	
QY	181 PDCQSLLRGMIEVDAAARLLTLEHIQHIWTIGGKNPEPEEQPIPRKVQIRLSLSLEDIDP 240
DB	
241 PDCQSLLRGMIEVDAAARLLTLEHIQHIWTIGGKNPEPEEQPIPRKVQIRLSLSLEDIDP 300	
QY	241 DVLDSMHSILGCFDRNKLQDLLSEENQEKMIYFLLLD RKERYPSQDEEDLP RNEIDP 300
DB	
301 DVLDSMHSILGCFDRNKLQDLLSEENQEKMIYFLLLD RKERYPSQDEEDLP RNEIDP 360	
QY	301 PRKRVDSPLMNRHGKRRPERKSMEVLSVT DGGSPVPARRAIEMAHQGRSRSISGASSGL 360
DB	
361 PRKRVDSPLMNRHGKRRPERKSMEVLSVT DGGSPVPARRAIEMAHQGRSRSISGASSGL 420	
QY	361 STSPLASPRVTPHPSPRGSLPTPKGTGPVHTPKESPAGTNP TTPSPSPSYGGVPMWRARLN 420
DB	
421 STSPLASPRVTPHPSPRGSLPTPKGTGPVHTPKESPAGTNP TTPSPSPSYGGVPMWRARLN 480	
QY	421 SIKNSFLGSPRFHRRKLQVTPPEMSNLTPESSPELAKS WFNGISLEKEEQIFVVIKD 480
DB	
481 SIKNSFLGSPRFHRRKLQVTPPEMSNLTPESSPELAKS WFNGISLEKEEQIFVVIKD 540	
QY	481 KPLSSIKADI VHAFLSTPSLSHVSISQTSFRAEYKATGGPAVTKQPKVQFVDITYTEGGE 540
DB	
541 KPLSSIKADI VHAFLSTPSLSHVSISQTSFRAEYKATGGPAVTKQPKVQFVDITYTEGGE 600	
QY	541 AQXENGYSYVTFLLSGPSRRFRKRVETIOAQLLSTHD PPAAOHLSDTTNCMEMMTGRLS 600
DB	
601 AQXENGYSYVTFLLSGPSRRFRKRVETIOAQLLSTHD PPAAOHLSDTTNCMEMMTGRLS 660	
QY	601 KCGIIPKS 608
DB	
661 KCGIIPKS 668	
RESULT 6	
US-10-283-247-2	
; Sequence 2, Application US/10283247	
; Publication No. US20030119037A1	
; GENERAL INFORMATION:	
; APPLICANT: NEELAM, Beena et al.	
; TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC	
; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES	
; TITLE OF INVENTION: THEREOF	
; FILE REFERENCE: CL001304	
; CURRENT APPLICATION NUMBER: US/10/283,247	
; CURRENT FILING DATE: 2002-10-30	
; NUMBER OF SEQ ID NOS: 10	
; SOFTWARE: FastSeq for Windows Version 4.0	
; SEQ ID NO 2	
; LENGTH: 674	
; TYPE: PRT	
; ORGANISM: Homo sapiens	
US-10-283-247-2	
Query Match 96.3%; Score 3071; DB 15; Length 674;	
Best Local Similarity 99.8%; Pred. No. 9e-189;	
Matches 586; Conservative 1; Mismatches 0; Indels 0; Gaps 0;	
QY	1 MKVEREIAILKLTIEHPHVLKLDHYVYENKKYLVLVLEHVS GGELFDYLVKKGRLLTPKEARK 60

	Db	61	MKVEREIAIKLIEHPHVLKHDVYENKKYLVLVLEHVSGGELFDLVVKGRUTPKARK	120
	Qy	61	FFRQIISALDFCHSHSICHRLDKPENLLDEKNIRIADFGMASLQVGDSLLETSCGSPH	120
	Db	121	FFRQIISALDFCHSHSICHRLDKPENLLDEKNIRIADFGMASLQVGDSLLETSCGSPH	180
	Qy	121	YACPEVIREKYDGRKADVWSCGVILFALLVGNLPEDDONLROLLEKVKRGVFHMPHFIP	180
	Db	181	YACPEVIREKYDGRKADVWSCGVILFALLVGNLPEDDONLROLLEKVKRGVFHMPHFIP	240
	Qy	181	PDQCSSLRGMIEWDAARRLTLEHIQKIHWIGCKNEPEPEQP1PRKVQIRSLPSLEDIDP	240
	Db	241	PDQCSSLRGMIEWDAARRLTLEHIQKIHWIGCKNEPEPEQP1PRKVQIRSLPSLEDIDP	300
	Qy	241	DVLDSMHSICGCFDRNRKLQDLLSEBENOEKMIFLLDRKERYPSQEDBDLPPRNEIDP	300
	Db	301	DVLDSMHSICGCFDRNRKLQDLLSEBENOEKMIFLLDRKERYPSQEDBDLPPRNEIDP	360
	Qy	301	PRKRVDSFMLNHRGKRPERKSMEVLSTVDGGSPVPARRAIEMAQHGORSISGASSGL	360
	Db	361	PRKRVDSFMLNHRGKRPERKSMEVLSTVDGGSPVPARRAIEMAQHGORSISGASSGL	420
	Qy	361	STSPSSPRVTPHSPRGSPLPTPKGTTPVHTPKESPAGTNPPTPPSPSVGGVPMRARLN	420
	Db	421	STSPSSPRVTPHSPRGSPLPTPKGTTPVHTPKESPAGTNPPTPPSPSVGGVPMRARLN	480
	Qy	421	SINKNFLGSRFRHRRKLOVPTPEMSNLTPESSPELAKSWFCNFISLEKEEQIFVVIKD	480
	Db	481	SINKNFLGSRFRHRRKLOVPTPEMSNLTPESSPELAKSWFCNFISLEKEEQIFVVIKD	540
	Qy	481	KPLSSIKADIIVHAFLSIPLSHSHSVISQTSFRAEYKATGGPAVFKPVKFQVDITYTEGGE	540
	Db	541	KPLSSIKADIIVHAFLSIPLSHSHSVISQTSFRAEYKATGGPAVFKPVKFQVDITYTEGGE	600
	Qy	541	AQENGIIYSVTFTLLSGPSRRFRKVVETITQAQLLSTHDPAAQHLS D 587	
	Db	601	AQENGIIYSVTFTLLSGPSRRFRKVVETITQAQLLSTHDPAAQHLS E 647	
 RESULT 7 US-10-283-247-9 ; Sequence 9, Application US/10283247 ; Publication No. US20030119037A1 ; GENERAL INFORMATION: ; APPLICANT: NEELAM, Beena et al. ; TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC ; ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES ; TITLE OF INVENTION: THEREOF ; FILE REFERENCE: CL001304 ; CURRENT APPLICATION NUMBER: US/10/283,247 ; CURRENT FILING DATE: 2002-10-30 ; NUMBER OF SEQ ID NOS: 10 ; SOFTWARE: FastSeq for Windows Version 4.0 ; SEQ ID NO 9 ; LENGTH: 614 ; TYPE: PRT ; ORGANISM: Homo sapiens US-10-283-247-9				
 Query Match 96.1%; Score 3065; DB 15; Length 614; Best Local Similarity 99.7%; Pred. No. 2e-188; Matches 589; Conservative 1; Mismatches 1; Indels 0; Gaps 0;				
	Qy	1	MKVEREIAIKLIEHPHVLKHDVYENKKYLVLVLEHVSGGELFDLVVKGRUTPKARK	60
	Db	1	MKVEREIAIKLIEHPHVLKHDVYENKKYLVLVLEHVSGGELFDLVVKGRUTPKARK	60
	Qy	61	FFRQIISALDFCHSHSICHRLDKPENLLDEKNIRIADFGMASLQVGDSLLETSCGSPH	120
	Db	61	FFRQIISALDFCHSHSICHRLDKPENLLDEKNIRIADFGMASLQVGDSLLETSCGSPH	120
	Qy	121	YACPEVIREKYDGRKADVWSCGVILFALLVGNLPEDDONLROLLEKVKRGVFHMPHFIP	180

Db 121 YACPEVIRGEKYGKADVWSCGVILFALLVGLPDDNLRQLLEKVRGVFHMHPHIP 180
QY 181 PDCQSLRGMIEVDAAARLTLEHIQKHIWYIGGKNEPEPEQIPRKVQIRSLPSLEDIDP 240
Db 181 PDCQSLRGMSEVDAAARLTLEHIQKHIWYIGGKNEPEPEQIPRKVQIRSLPSLEDIDP 240
QY 241 DVLDSMHSGLCFDRDNKLLQDLLSEENQEKMIYFLLDRKERYPSQDEDLPPRNEIDP 300
Db 241 DVLDSMHSGLCFDRDNKLLQDLLSEENQEKMIYFLLDRKERYPSQDEDLPPRNEIDP 300
QY 301 PRKRVDSFPMNRHGKRRPERKSMELVSVTDGSPVPARRAEMAQHGORSISGASSGL 360
Db 301 PRKRVDSFPMNRHGKRRPERKSMELVSVTDGSPVPARRAEMAQHGORSISGASSGL 360
QY 361 STPSLSSPRVTPHPSRPGSPPLTPKGTVPHTPKESPAGTNPPTPPSSPSVGVWPWRARLN 420
Db 361 STPSLSSPRVTPHPSRPGSPPLTPKGTVPHTPKESPAGTNPPTPPSSPSVGVWPWRARLN 420
QY 421 SIKNSFLGSPFRHRRKLOVPTPEEMSNLTPESSPELAKSWFGNFISLEKEEQIFVVIKD 480
Db 421 SIKNSFLGSPFRHRRKLOVPTPEEMSNLTPESSPELAKSWFGNFISLEKEEQIFVVIKD 480
QY 481 KPLSSIKADIHAFSLIPSLSHSVISQTSFRAEYKATGCPAVFQKPVKFOVDITTYEGGE 540
Db 481 KPLSSIKADIHAFSLIPSLSHSVISQTSFRAEYKATGCPAVFQKPVKFOVDITTYEGGE 540
QY 541 AQKENGIVSVTFTLLSGPSRRFRKRVVETIQALLSTHDPAPAAQLHSD 587
Db 541 AQKENGIVSVTFTLLSGPSRRFRKRVVETIQALLSTHDPAPAAQLHSD 587

RESULT 8
US-10-283-247-10
; Sequence 10, Application US/10283247
; Publication No. US20030119037A1
; GENERAL INFORMATION:
; APPLICANT: NEELAM, Beena et al.
; TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC
; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES
; FILE REFERENCE: CL001304
; CURRENT APPLICATION NUMBER: US/10/283,247
; CURRENT FILING DATE: 2002-10-30
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 10
; LENGTH: 614
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-283-247-10

Query Match 96.1%; Score 3065; DB 15; Length 614;
Best Local Similarity 99.7%; Pred. No. 2e-188;
Matches 585; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
QY 1 MKVEREIAILKLIIEHPHVLKLDVYENKYLVLVLEHVSQGGELFDYLVKKGLRTPKEARK 60
Db 1 MKVEREIAILKLIIEHPHVLKLDVYENKYLVLVLEHVSQGGELFDYLVKKGLRTPKEARK 60
QY 61 PFRQIISALDFCHSHSICHRLDKPENLLDEKNNIRIADFGMASLQVGDLSLETSCGSPH 120
Db 61 PFRQIISALDFCHSHSICHRLDKPENLLDEKNNIRIADFGMASLQVGDLSLETSCGSPH 120
QY 121 YACPEVIRGEKYGKADVWSCGVILFALLVGLPDDNLRQLLEKVRGVFHMHPHIP 180
Db 121 YACPEVIRGEKYGKADVWSCGVILFALLVGLPDDNLRQLLEKVRGVFHMHPHIP 180
QY 181 PDCQSLRGMIEVDAAARLTLEHIQKHIWYIGGKNEPEPEQIPRKVQIRSLPSLEDIDP 240
Db 181 PDCQSLRGMSEVDAAARLTLEHIQKHIWYIGGKNEPEPEQIPRKVQIRSLPSLEDIDP 240
QY 241 DVLDSMHSGLCFDRDNKLLQDLLSEENQEKMIYFLLDRKERYPSQDEDLPPRNEIDP 300
Db 241 DVLDSMHSGLCFDRDNKLLQDLLSEENQEKMIYFLLDRKERYPSQDEDLPPRNEIDP 300

Db 241 DVLDSMHSGLCFDRDNKLLQDLLSEENQEKMIYFLLDRKERYPSQDEDLPPRNEIDP 300
QY 301 PRKRVDSFPMNRHGKRRPERKSMELVSVTDGSPVPARRAEMAQHGORSISGASSGL 360
Db 301 PRKRVDSFPMNRHGKRRPERKSMELVSVTDGSPVPARRAEMAQHGORSISGASSGL 360
QY 361 STPSLSSPRVTPHPSRPGSPPLTPKGTVPHTPKESPAGTNPPTPPSSPSVGVWPWRARLN 420
Db 361 STPSLSSPRVTPHPSRPGSPPLTPKGTVPHTPKESPAGTNPPTPPSSPSVGVWPWRARLN 420
QY 421 SIKNSFLGSPFRHRRKLOVPTPEEMSNLTPESSPELAKSWFGNFISLEKEEQIFVVIKD 480
Db 421 SIKNSFLGSPFRHRRKLOVPTPEEMSNLTPESSPELAKSWFGNFISLEKEEQIFVVIKD 480
QY 481 KPLSSIKADIHAFSLIPSLSHSVISQTSFRAEYKATGCPAVFQKPVKFOVDITTYEGGE 540
Db 481 KPLSSIKADIHAFSLIPSLSHSVISQTSFRAEYKATGCPAVFQKPVKFOVDITTYEGGE 540
QY 541 AQKENGIVSVTFTLLSGPSRRFRKRVVETIQALLSTHDPAPAAQLHSD 587
Db 541 AQKENGIVSVTFTLLSGPSRRFRKRVVETIQALLSTHDPAPAAQLHSD 587

RESULT 9
US-09-842-582-2
; Sequence 2, Application US/09842582
; Patent No. US20020155570A1
; GENERAL INFORMATION:
; APPLICANT: Millennium Pharmaceuticals, Inc.
; APPLICANT: Meyers, Rachel
; TITLE OF INVENTION: 2246, NOVEL PROTEIN KINASE MOLECULES AND
; TITLE OF INVENTION: USES THEREFOR
; FILE REFERENCE: 38155-20054.00
; CURRENT APPLICATION NUMBER: US/09/842,582
; CURRENT FILING DATE: 2001-04-25
; PRIOR FILING DATE: 2000-04-25
; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 674
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-842-582-2

Query Match 96.1%; Score 3065; DB 10; Length 674;
Best Local Similarity 99.7%; Pred. No. 2.2e-188;
Matches 585; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
QY 1 MKVEREIAILKLIIEHPHVLKLDVYENKYLVLVLEHVSQGGELFDYLVKKGLRTPKEARK 60
Db 61 MKVEREIAILKLIIEHPHVLKLDVYENKYLVLVLEHVSQGGELFDYLVKKGLRTPKEARK 120
QY 61 PFRQIISALDFCHSHSICHRLDKPENLLDEKNNIRIADFGMASLQVGDLSLETSCGSPH 120
Db 121 PFRQIISALDFCHSHSICHRLDKPENLLDEKNNIRIADFGMASLQVGDLSLETSCGSPH 180
QY 121 YACPEVIRGEKYGKADVWSCGVILFALLVGLPDDNLRQLLEKVRGVFHMHPHIP 180
Db 181 YACPEVIRGEKYGKADVWSCGVILFALLVGLPDDNLRQLLEKVRGVFHMHPHIP 240
QY 181 PDCQSLRGMIEVDAAARLTLEHIQKHIWYIGGKNEPEPEQIPRKVQIRSLPSLEDIDP 240
Db 241 PDCQSLRGMSEVDAAARLTLEHIQKHIWYIGGKNEPEPEQIPRKVQIRSLPSLEDIDP 300
QY 241 DVLDSMHSGLCFDRDNKLLQDLLSEENQEKMIYFLLDRKERYPSQDEDLPPRNEIDP 300
Db 301 DVLDSMHSGLCFDRDNKLLQDLLSEENQEKMIYFLLDRKERYPSQDEDLPPRNEIDP 360
QY 301 PRKRVDSFPMNRHGKRRPERKSMELVSVTDGSPVPARRAEMAQHGORSISGASSGL 360
Db 361 PRKRVDSFPMNRHGKRRPERKSMELVSVTDGSPVPARRAEMAQHGORSISGASSGL 420


```
QY 361 STSPLSSPRVTPHSPRGSPLTPKGTPTVHTPKESPAGTNPPTPPSPSVGVPWRARLN 420
DB 421 STSPLSSPRVTPHSPRGSPLTPKGTPTVHTPKESPAGTNPPTPPSPSVGVPWRARLN 480
QY 421 SIKNSFLGSPRHRRLKQVPTPEMSNLTPESSPELAKSWFGNFI SLEKEEQIFVWIKD 480
DB 481 SIKNSFLGSPRHRRLKQVPTPEMSNLTPESSPELAKSWFGNFI SLEKEEQIFVWIKD 540
QY 481 KPLSSIKADIHAFSLISPSLSHSHVISOTSFAEYKATGGPAVFQKPVKFOVDIITYTEGGE 540
DB 541 KPLSSIKADIHAFSLISPSLSHSHVISOTSFAEYKATGGPAVFQKPVKFOVDIITYTEGGE 600
QY 541 AQKENGISYVFTLLSGSPRRFRKRVVETIQALLSTHDPAPAAQHLSD 587
DB 601 AQKENGISYVFTLLSGSPRRFRKRVVETIQALLSTHDPAPAAQHLSE 647

RESULT 10
US-10-283-247-7
; Sequence 7, Application US/10283247
; Publication No. US20030119037A1
; GENERAL INFORMATION:
; APPLICANT: NEELAM, Beena et al.
; TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC
; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES
; FILE REFERENCE: CL001304
; CURRENT APPLICATION NUMBER: US/10/283,247
; CURRENT FILING DATE: 2002-10-30
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 7
; LENGTH: 674
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-283-247-7

Query Match 96.1%; Score 3065; DB 15; Length 674;
Best Local Similarity 99.7%; Pred. No. 2.2e-188;
Matches 585; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 MKVEREIAILKLIIEHPHVLKLDHVVYENKKYLYLVLEHVSQGGELFDYLVKKGLRTPKEARK 60
DB 61 MKVEREIAILKLIIEHPHVLKLDHVVYENKKYLYLVLEHVSQGGELFDYLVKKGLRTPKEARK 120
QY 61 FFRQIISALDFCHSHSICHRLDKPENLLDEKNNIRIADFGMASLQVSDLSLETSCGSPH 120
DB 121 FFRQIISALDFCHSHSICHRLDKPENLLDEKNNIRIADFGMASLQVSDLSLETSCGSPH 180
QY 121 YACPEVIRGEKYDGRKADVWSCGVILFALLVGLPFDNRLQLLEKVKRGVFMHPHIP 180
DB 181 YACPEVIRGEKYDGRKADVWSCGVILFALLVGLPFDNRLQLLEKVKRGVFMHPHIP 240
QY 181 PDCOSLLRGMIENVDAARLTLEHIOKHIWIIGKNEPEPEQIPRKVQIRSLPSLEDIDP 240
DB 241 PDCOSLLRGMIENVDAARLTLEHIOKHIWIIGKNEPEPEQIPRKVQIRSLPSLEDIDP 300
QY 241 YACPEVIRGEKYDGRKADVWSCGVILFALLVGLPFDNRLQLLEKVKRGVFMHPHIP 180
DB 181 YACPEVIRGEKYDGRKADVWSCGVILFALLVGLPFDNRLQLLEKVKRGVFMHPHIP 240
QY 181 PDCOSLLRGMIENVDAARLTLEHIOKHIWIIGKNEPEPEQIPRKVQIRSLPSLEDIDP 240
DB 241 PDCOSLLRGMIENVDAARLTLEHIOKHIWIIGKNEPEPEQIPRKVQIRSLPSLEDIDP 300
QY 241 DVLDMSHSLGCFDRDNKLLQDLLSEENQEKMIYFLLDLDRKERYPSQDEDLPPRNEIDP 300
DB 301 DVLDMSHSLGCFDRDNKLLQDLLSEENQEKMIYFLLDLDRKERYPSQDEDLPPRNEIDP 360
QY 301 PRKRVDSFPLNRHGKRRPERKSMELSVTDGSGSPVARRAIEAQAHQGRSISGASSGL 360
DB 361 PRKRVDSFPLNRHGKRRPERKSMELSVTDGSGSPVARRAIEAQAHQGRSISGASSGL 420
QY 361 STSPLSSPRVTPHSPRGSPLTPKGTPTVHTPKESPAGTNPPTPPSPSVGVPWRARLN 420
DB 421 STSPLSSPRVTPHSPRGSPLTPKGTPTVHTPKESPAGTNPPTPPSPSVGVPWRARLN 480
QY 421 SIKNSFLGSPRHRRLKQVPTPEMSNLTPESSPELAKSWFGNFI SLEKEEQIFVWIKD 480
DB 481 SIKNSFLGSPRHRRLKQVPTPEMSNLTPESSPELAKSWFGNFI SLEKEEQIFVWIKD 540
QY 481 KPLSSIKADIHAFSLISPSLSHSHVISOTSFAEYKATGGPAVFQKPVKFOVDIITYTEGGE 540
DB 541 KPLSSIKADIHAFSLISPSLSHSHVISOTSFAEYKATGGPAVFQKPVKFOVDIITYTEGGE 600
QY 541 AQKENGISYVFTLLSGSPRRFRKRVVETIQALLSTHDPAPAAQHLSD 587
DB 601 AQKENGISYVFTLLSGSPRRFRKRVVETIQALLSTHDPAPAAQHLSE 647
```

```
QY 481 KPLSSIKADIHAFSLISPSLSHSHVISOTSFAEYKATGGPAVFQKPVKFOVDIITYTEGGE 540
DB 541 KPLSSIKADIHAFSLISPSLSHSHVISOTSFAEYKATGGPAVFQKPVKFOVDIITYTEGGE 600
QY 541 AQKENGISYVFTLLSGSPRRFRKRVVETIQALLSTHDPAPAAQHLSD 587
DB 601 AQKENGISYVFTLLSGSPRRFRKRVVETIQALLSTHDPAPAAQHLSE 647

RESULT 11
US-10-283-247-8
; Sequence 8, Application US/10283247
; Publication No. US20030119037A1
; GENERAL INFORMATION:
; APPLICANT: NEELAM, Beena et al.
; TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC
; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES
; FILE REFERENCE: CL001304
; CURRENT APPLICATION NUMBER: US/10/283,247
; CURRENT FILING DATE: 2002-10-30
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 8
; LENGTH: 674
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-283-247-8

Query Match 96.1%; Score 3065; DB 15; Length 674;
Best Local Similarity 99.7%; Pred. No. 2.2e-188;
Matches 585; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 MKVEREIAILKLIIEHPHVLKLDHVVYENKKYLYLVLEHVSQGGELFDYLVKKGLRTPKEARK 60
DB 61 MKVEREIAILKLIIEHPHVLKLDHVVYENKKYLYLVLEHVSQGGELFDYLVKKGLRTPKEARK 120
QY 61 FFRQIISALDFCHSHSICHRLDKPENLLDEKNNIRIADFGMASLQVSDLSLETSCGSPH 120
DB 121 FFRQIISALDFCHSHSICHRLDKPENLLDEKNNIRIADFGMASLQVSDLSLETSCGSPH 180
QY 121 YACPEVIRGEKYDGRKADVWSCGVILFALLVGLPFDNRLQLLEKVKRGVFMHPHIP 180
DB 181 YACPEVIRGEKYDGRKADVWSCGVILFALLVGLPFDNRLQLLEKVKRGVFMHPHIP 240
QY 181 PDCOSLLRGMIENVDAARLTLEHIOKHIWIIGKNEPEPEQIPRKVQIRSLPSLEDIDP 240
DB 241 PDCOSLLRGMIENVDAARLTLEHIOKHIWIIGKNEPEPEQIPRKVQIRSLPSLEDIDP 300
QY 241 DVLDMSHSLGCFDRDNKLLQDLLSEENQEKMIYFLLDLDRKERYPSQDEDLPPRNEIDP 300
DB 301 DVLDMSHSLGCFDRDNKLLQDLLSEENQEKMIYFLLDLDRKERYPSQDEDLPPRNEIDP 360
QY 301 PRKRVDSFPLNRHGKRRPERKSMELSVTDGSGSPVARRAIEAQAHQGRSISGASSGL 360
DB 361 PRKRVDSFPLNRHGKRRPERKSMELSVTDGSGSPVARRAIEAQAHQGRSISGASSGL 420
QY 361 STSPLSSPRVTPHSPRGSPLTPKGTPTVHTPKESPAGTNPPTPPSPSVGVPWRARLN 420
DB 421 STSPLSSPRVTPHSPRGSPLTPKGTPTVHTPKESPAGTNPPTPPSPSVGVPWRARLN 480
QY 421 SIKNSFLGSPRHRRLKQVPTPEMSNLTPESSPELAKSWFGNFI SLEKEEQIFVWIKD 480
DB 481 SIKNSFLGSPRHRRLKQVPTPEMSNLTPESSPELAKSWFGNFI SLEKEEQIFVWIKD 540
QY 481 KPLSSIKADIHAFSLISPSLSHSHVISOTSFAEYKATGGPAVFQKPVKFOVDIITYTEGGE 540
DB 541 KPLSSIKADIHAFSLISPSLSHSHVISOTSFAEYKATGGPAVFQKPVKFOVDIITYTEGGE 600
QY 541 AQKENGISYVFTLLSGSPRRFRKRVVETIQALLSTHDPAPAAQHLSD 587
DB 601 AQKENGISYVFTLLSGSPRRFRKRVVETIQALLSTHDPAPAAQHLSE 647
```

RESULT 12

US-10-283-247-5
; Sequence 5, Application US/10283247
; Publication No. US20030119037A1
; GENERAL INFORMATION:
; APPLICANT: NEELAM, Beena et al.
; TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC
; ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES
; TITLE OF INVENTION: THEREOF
; FILE REFERENCE: CL001304
; CURRENT APPLICATION NUMBER: US/10/283,247
; CURRENT FILING DATE: 2002-10-30
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 5
; LENGTH: 636
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-283-247-5

Query Match 95.7%; Score 3050; DB 15; Length 636;
Best Local Similarity 96.2%; Pred. No. 1.9e-187;
Matches 586; Conservative 1; Mismatches 0; Indels 22; Gaps 1;

Qy 1 MKVEREIAILKLIIEHPHVLKLDVYENKKYLVLVLEHVSGGELFDYLVKGRITPKARK 60
Db 1 MKVEREIAILKLIIEHPHVLKLDVYENKKYLVLVLEHVSGGELFDYLVKGRITPKARK 60

Qy 61 FFRQIISALDFCHSHSICHRLKPNLLEDEKNNIRIADFGWASLQVDSLETSCGSPH 120
Db 61 FFRQIISALDFCHSHSICHRLKPNLLEDEKNNIRIADFGWASLQVDSLETSCGSPH 120

Qy 121 YACPEVIRGEKYDGRKADWVSCGVTILFALLGALPFDNRLQLLEKVRGVFMPHFIP 180
Db 121 YACPEVIRGEKYDGRKADWVSCGVTILFALLGALPFDNRLQLLEKVRGVFMPHFIP 180

Qy 181 PDCOSLLRGMIEVDAARLTLEHQQHVIWYIGGKNEPEPEPIPRKQVIRSLPSLEIDP 240
Db 181 PDCOSLLRGMIEVDAARLTLEHQQHVIWYIGGKNEPEPEPIPRKQVIRSLPSLEIDP 240

Qy 241 DVLDSMHSGLCFDRNKLQDLLEENQEKMIYFLLDLRKERYPSQDEDELPNRNEIDP 300
Db 241 DVLDSMHSGLCFDRNKLQDLLEENQEKMIYFLLDLRKERYPSQDEDELPNRNEIDP 300

Qy 301 PRKRVDSPLNRHGKRRPERKSMELSVTDGSGSPVARRAITEMAQHGO----- 348
Db 301 PRKRVDSPLNRHGKRRPERKSMELSVTDGSGSPVARRAITEMAQHGO----- 348

Qy 349 -----RSRSISGASSGLSTSPSSPRVTPHPSPRGSPPLTPKGTPTVHTPKESPAG 398
Db 361 EAHQFQSKEDRSRISGASSGLSTSPSSPRVTPHPSPRGSPPLTPKGTPTVHTPKESPAG 420

Qy 399 TPNPTPSSPSVGVVPMARLNSIKNSFLGSPRFRHRLKQVPTPEMSNLTPSSPELAK 458
Db 421 TPNPTPSSPSVGVVPMARLNSIKNSFLGSPRFRHRLKQVPTPEMSNLTPSSPELAK 480

Qy 459 KSWFGNFTLSLEKEQIFVVIKDKPLSSIKADIVHAFSLIPSLSHSVISQTSFRAYKATG 518
Db 481 KSWFGNFTLSLEKEQIFVVIKDKPLSSIKADIVHAFSLIPSLSHSVISQTSFRAYKATG 540

Qy 519 GPAVFQKPVKFQVDITYTTEGGEAKENGIVSVTFTLLSGPSRRFRKRVETIQALLSTHD 578
Db 541 GPAVFQKPVKFQVDITYTTEGGEAKENGIVSVTFTLLSGPSRRFRKRVETIQALLSTHD 600

Qy 579 PPAHQHLS 587
Db 601 PPAHQHLS 609

RESULT 13

US-10-195-072-4

; Sequence 4, Application US/10195072
; Publication No. US20030092036A1
; GENERAL INFORMATION:
; APPLICANT: Origene Technologies
; TITLE OF INVENTION: Full-Length Serine Protein Kinase in Brain and Pancreas
; FILE REFERENCE: 16U 101 C2
; CURRENT APPLICATION NUMBER: US/10/195,072
; CURRENT FILING DATE: 2002-07-15
; PRIOR APPLICATION NUMBER: US 09/930,181
; PRIOR FILING DATE: 2001-08-16
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 4
; LENGTH: 585
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-195-072-4

Query Match 95.0%; Score 3028; DB 15; Length 585;
Best Local Similarity 100.0%; Pred. No. 4.3e-186;
Matches 577; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 32 YLVLEHVSGGELFDYLVKGRITPKARKFFROIISALDFCHSHSICHRLKPNLLEDE 91
Db 9 YLVLEHVSGGELFDYLVKGRITPKARKFFROIISALDFCHSHSICHRLKPNLLEDE 68

Qy 92 KNNIRIADFGWASLQVDSLETSCGSPHYACPEVIRGEKYDGRKADWVSCGVTILFALLV 151
Db 69 KNNIRIADFGWASLQVDSLETSCGSPHYACPEVIRGEKYDGRKADWVSCGVTILFALLV 128

Qy 152 GALPFDNRLQLLEKVRGVFMPHFIPDPCOSLLRGMIEVDAARLTLEHQQHVIWYI 211
Db 129 GALPFDNRLQLLEKVRGVFMPHFIPDPCOSLLRGMIEVDAARLTLEHQQHVIWYI 188

Qy 212 GGNKEPEPEPIPRKQVIRSLPSLEIDPVLDSMHSGLCFDRNKLQDLLEENQEK 271
Db 189 GGNKEPEPEPIPRKQVIRSLPSLEIDPVLDSMHSGLCFDRNKLQDLLEENQEK 248

Qy 272 MIYFLLDLRKERYPSQDEDELPNRNEIDPDKRVDSPLNRHGKRRPERKSMELSVTDG 331
Db 249 MIYFLLDLRKERYPSQDEDELPNRNEIDPDKRVDSPLNRHGKRRPERKSMELSVTDG 308

Qy 332 GSPVARRAITEMAQHGORSRISGASSGLSTSPSSPRVTPHPSPRGSPPLTPKGTPTVHT 391
Db 309 GSPVARRAITEMAQHGORSRISGASSGLSTSPSSPRVTPHPSPRGSPPLTPKGTPTVHT 368

Qy 392 PKESPAGTPTNPTPSSPSVGVVPMARLNSIKNSFLGSPRFRHRLKQVPTPEMSNLTP 451
Db 369 PKESPAGTPTNPTPSSPSVGVVPMARLNSIKNSFLGSPRFRHRLKQVPTPEMSNLTP 428

Qy 452 SSPELAKSWFGNFTLSLEKEQIFVVIKDKPLSSIKADIVHAFSLIPSLSHSVISQTSFR 511
Db 429 SSPELAKSWFGNFTLSLEKEQIFVVIKDKPLSSIKADIVHAFSLIPSLSHSVISQTSFR 488

Qy 512 AYEKATGGPAVFQKPVKFQVDITYTTEGGEAKENGIVSVTFTLLSGPSRRFRKRVETIQ 571
Db 489 AYEKATGGPAVFQKPVKFQVDITYTTEGGEAKENGIVSVTFTLLSGPSRRFRKRVETIQ 548

Qy 572 QLLSTHDPAAQHLSDTTNCMMMTGRLSKCGIIPKS 608
Db 549 QLLSTHDPAAQHLSDTTNCMMMTGRLSKCGIIPKS 595

RESULT 14

US-10-195-071-4
; Sequence 4, Application US/10195071
; Publication No. US20030096271A1
; GENERAL INFORMATION:
; APPLICANT: Origene Technologies
; TITLE OF INVENTION: Full-Length Serine Protein Kinase in Brain and Pancreas
; FILE REFERENCE: 16U 101 C1
; CURRENT APPLICATION NUMBER: US/10/195,071
; CURRENT FILING DATE: 2002-07-15

;; PRIOR APPLICATION NUMBER: US 09/930,181
;; PRIOR FILING DATE: 2001-08-16
;; NUMBER OF SEQ ID NOS: 18
;; SOFTWARE: PatentIn version 3.1
;; SEQ ID NO 4
;; LENGTH: 585
;; TYPE: PRT
;; ORGANISM: Homo sapiens
US-10-195-072-4

Query Match 95.0%; Score 3028; DB 15; Length 585;
Best Local Similarity 100.0%; Pred. No. 4.3e-186;
Matches 577; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 32 YLVLEHVSQGGELFYLVKKGRLTPKEARKFQIISALDCHSHSICHRLKPENLLDE 91
Db 9 YLVLEHVSQGGELFYLVKKGRLTPKEARKFQIISALDCHSHSICHRLKPENLLDE 68

Qy 92 KNNIRIADFGWASIQVGSLLTSCGSPHYACPEVIRGEKYDGRKADVWSCGVILFALLV 151
Db 69 KNNIRIADFGWASIQVGSLLTSCGSPHYACPEVIRGEKYDGRKADVWSCGVILFALLV 128

Qy 152 GALPFDNDNLQQLLEKVKRGVFMHPHFIPDPQCSLLRGMIEVDAAARLTLEHIQKHIWI 211
Db 129 GALPFDNDNLQQLLEKVKRGVFMHPHFIPDPQCSLLRGMIEVDAAARLTLEHIQKHIWI 108

Qy 212 GKKNEPEEQIPRKVQIRSLPSLEDIDPDVLDMSHSLGCFDRNKLQDLISEENOEK 271
Db 189 GKKNEPEEQIPRKVQIRSLPSLEDIDPDVLDMSHSLGCFDRNKLQDLISEENOEK 248

Qy 272 MIYFLLDRKERYPSQEDIDLPRNEIDPPKRVDSPLNHRGKRRPERKSMEVLSVTDG 331
Db 249 MIYFLLDRKERYPSQEDIDLPRNEIDPPKRVDSPLNHRGKRRPERKSMEVLSVTDG 308

Qy 332 GSPVARRAIEMAHQGRSISGASGLSTSPSSPRVTPHPSRGSPLTPKGTVPHT 391
Db 309 GSPVARRAIEMAHQGRSISGASGLSTSPSSPRVTPHPSRGSPLTPKGTVPHT 368

Qy 392 PKESPAGTNPPTSSPSVGVVWRRARLNSIKNSFLGSPRHRKLQVPTPEEMSNLTPE 451
Db 369 PKESPAGTNPPTSSPSVGVVWRRARLNSIKNSFLGSPRHRKLQVPTPEEMSNLTPE 428

Qy 452 SSPELAKKSWFGNFISLEKEQIFVWIKDKPLSSIKADIVHAFSLSPSHSVISQTSFR 511
Db 429 SSPELAKKSWFGNFISLEKEQIFVWIKDKPLSSIKADIVHAFSLSPSHSVISQTSFR 488

Qy 512 AERYKATGPVAFQKPVKQVDITYTEGGEAOKENGISYVTFLLSGPSRRRFRKRVVETIOA 571
Db 489 AERYKATGPVAFQKPVKQVDITYTEGGEAOKENGISYVTFLLSGPSRRRFRKRVVETIOA 548

Qy 572 QLLSTHDPAAQHLSDTTNCMEMMTGRLSKGIIPKS 608
Db 549 QLLSTHDPAAQHLSDTTNCMEMMTGRLSKGIIPKS 585

RESULT 15
US-10-195-072-17
; Sequence 17, Application US/10195072
; Publication No. US20030092036A1
; GENERAL INFORMATION:
; APPLICANT: Origene Technologies
; TITLE OF INVENTION: Full-Length Serine Protein Kinase in Brain and Pancreas
; FILE REFERENCE: 16U 101 C2
; CURRENT APPLICATION NUMBER: US/10/195,072
; CURRENT FILING DATE: 2002-07-15
; PRIOR APPLICATION NUMBER: US 09/930,181
; PRIOR FILING DATE: 2001-08-16
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 17
; LENGTH: 603
; TYPE: PRT
; ORGANISM: Homo sapiens

US-10-195-072-17

Query Match 94.6%; Score 3015; DB 15; Length 603;
Best Local Similarity 99.7%; Pred. No. 3.1e-185;
Matches 574; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 12 LIEHPHVLKLDVYENKYLVLVLEHVSQGGELFYLVKKGRLTPKEARKFQIISALDF 71
Db 1 LIEHPHVLKLDVYENKYLVLVLEHVSQGGELFYLVKKGRLTPKEARKFQIISALDF 60

Qy 72 CHSHSICHRLKPENLLDEKNNIRIADFGWASIQVGSLLTSCGSPHYACPEVIRGEK 131
Db 61 CHSHSICHRLKPENLLDEKNNIRIADFGWASIQVGSLLTSCGSPHYACPEVIRGEK 120

Qy 132 YDGRKADVWSCGVILFALLVWGLPFDNDNLQQLLEKVKRGVFMHPHFIPDPQCSLLRGM 191
Db 121 YDGRKADVWSCGVILFALLVWGLPFDNDNLQQLLEKVKRGVFMHPHFIPDPQCSLLRGM 180

Qy 192 EYDAAARLTLEHIQKHIWIYIGGKNEPEEQIPRKVQIRSLPSLEDIDPDVLDMSHSLG 251
Db 181 EYDAAARLTLEHIQKHIWIYIGGKNEPEEQIPRKVQIRSLPSLEDIDPDVLDMSHSLG 240

Qy 252 FRDNKLQDLISEENOEKMIYFLLDRKERYPSQEDIDLPRNEIDPPKRVDSPLMLN 311
Db 241 FRDNKLQDLISEENOEKMIYFLLDRKERYPSQEDIDLPRNEIDPPKRVDSPLMLN 300

Qy 312 RHGKRPERKSMEVLSVTDGSPVARRAIEMAHQGRSISGASGLSTSPSSPRVT 371
Db 301 RHGKRPERKSMEVLSVTDGSPVARRAIEMAHQGRSISGASGLSTSPSSPRVT 360

Qy 372 PHPSRGSPLTPKGTVPHTPKESPAGTNPPTSSPSVGVVWRRARLNSIKNSFLGSPR 431
Db 361 PHPSRGSPLTPKGTVPHTPKESPAGTNPPTSSPSVGVVWRRARLNSIKNSFLGSPR 420

Qy 432 FHRKQLQVPTPEEMSNLTPESSPELAKKSWFGNFISLEKEQIFVWIKDKPLSSIKADIV 491
Db 421 FHRKQLQVPTPEEMSNLTPESSPELAKKSWFGNFISLEKEQIFVWIKDKPLSSIKADIV 480

Qy 492 HAFSLSPSHSVISQTSFRAEYKATGGPAVFOKPVKQVDITYTEGGEAOKENGISYVT 551
Db 481 HAFSLSPSHSVISQTSFRAEYKATGGPAVFOKPVKQVDITYTEGGEAOKENGISYVT 540

Qy 552 FTLLSGPSRRRFRKRVVETIOAQLLSTHDPAAQHLSD 587
Db 541 FTLLSGPSRRRFRKRVVETIOAQLLSTHDPAAQHLSE 576

Search completed: November 26, 2003, 12:11:03
Job time : 27.7304 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 26, 2003, 11:58:31 ; Search time 15.7053 Seconds
(without alignments)
1799.622 Million cell updates/sec

Title: US-10-054-579-2

Perfect score: 3497

Sequence: 1 MTSTGDKGGAQYVGPYR.....TNCMEMTGRSLKCGIIPKS 668

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 328717 seqs, 42310858 residues

Total number of hits satisfying chosen parameters: 328717

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents AA.*

- 1: /cgn2_6/ptodata/1/iaa/5A.COMB.pap.*
- 2: /cgn2_6/ptodata/1/iaa/5B.COMB.pap.*
- 3: /cgn2_6/ptodata/1/iaa/6A.COMB.pap.*
- 4: /cgn2_6/ptodata/1/iaa/6B.COMB.pap.*
- 5: /cgn2_6/ptodata/1/iaa/ECTUS.COMB.pap.*
- 6: /cgn2_6/ptodata/1/iaa/backfiles.pap.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	3497	100.0	668	US-09-930-181-2	Sequence 2, Appli
2	3028	86.6	585	US-09-930-181-4	Sequence 4, Appli
3	3015	86.2	603	US-09-930-181-17	Sequence 17, Appl
4	796.5	22.8	745	US-09-523-849-36	Sequence 36, Appl
5	793.5	22.7	724	US-09-984-890-2	Sequence 2, Appli
6	792.5	22.7	722	US-09-984-890-4	Sequence 4, Appli
7	790.5	22.6	722	US-08-817-8328-32	Sequence 32, Appl
8	768.5	22.0	729	US-08-677-298-2	Sequence 2, Appli
9	768.5	22.0	729	US-09-523-849-33	Sequence 33, Appl
10	764.5	21.9	793	US-09-523-849-32	Sequence 32, Appl
11	749	21.4	776	US-09-523-849-34	Sequence 34, Appl
12	729.5	20.9	552	US-08-817-8328-31	Sequence 31, Appl
13	720.5	20.6	633	US-08-557-006C-40	Sequence 40, Appl
14	718.5	20.5	345	US-08-557-006C-43	Sequence 43, Appl
15	694.5	19.9	257	US-09-101-146-1	Sequence 1, Appli
16	681	19.5	257	US-07-857-224B-6	Sequence 6, Appli
17	680	19.4	149	US-09-930-181-18	Sequence 18, Appl
18	670.5	19.2	604	US-09-523-849-35	Sequence 35, Appl
19	599.5	17.1	631	US-09-579-664B-11	Sequence 11, Appl
20	523	15.0	290	US-09-734-673-4	Sequence 4, Appli
21	523	15.0	334	US-09-523-849-31	Sequence 31, Appl
22	520.5	14.9	252	US-07-857-224B-26	Sequence 26, Appl
23	515.5	14.7	436	US-09-734-673-2	Sequence 2, Appli
24	515.5	14.7	436	US-09-523-849-2	Sequence 2, Appli
25	506.5	14.5	260	US-07-857-224B-27	Sequence 27, Appl
26	501	14.3	353	US-08-688-988-31	Sequence 31, Appl

28	491.5	14.1	556	4	US-09-800-960-4	Sequence 4, Appli
29	490	14.0	363	3	US-08-688-988-30	Sequence 30, Appl
30	488.5	14.0	260	3	US-07-857-224B-28	Sequence 28, Appl
31	485	13.9	339	3	US-08-688-988-13	Sequence 33, Appl
32	485	13.9	351	3	US-08-688-988-28	Sequence 28, Appl
33	485	13.9	565	4	US-09-800-960-2	Sequence 2, Appli
34	481	13.8	354	3	US-08-688-988-29	Sequence 29, Appl
35	471.5	13.5	359	3	US-08-688-988-32	Sequence 32, Appl
36	469.5	13.4	1037	4	US-09-428-711A-21	Sequence 21, Appl
37	467.5	13.4	339	3	US-08-688-988-2	Sequence 2, Appli
38	466.5	13.3	295	1	US-07-951-715A-23	Sequence 23, Appl
39	466.5	13.3	295	2	US-08-459-448A-23	Sequence 23, Appl
40	466.5	13.3	295	3	US-08-459-595A-23	Sequence 23, Appl
41	466.5	13.3	295	3	US-08-459-504B-23	Sequence 23, Appl
42	466.5	13.3	295	3	US-08-459-44A-23	Sequence 23, Appl
43	466.5	13.3	295	4	US-09-547-422-23	Sequence 23, Appl
44	463.5	13.3	433	2	US-08-913-050A-7	Sequence 7, Appli
45	463.5	13.3	433	2	US-08-749-902-5	Sequence 5, Appli

ALIGNMENTS

RESULT 1

US-09-930-181-2

Sequence 2, Application US/09930181

Patent No. 6455292

GENERAL INFORMATION:

APPLICANT: Origene Technologies

TITLE OF INVENTION: Full-Length Serine Protein Kinase in Brain and Pancreas

FILE REFERENCE: 160 101 V1

CURRENT APPLICATION NUMBER: US/09/930,181

CURRENT FILING DATE: 2001-08-16

NUMBER OF SEQ ID NOS: 18

SOFTWARE: PatentIn version 3.0

SEQ ID NO 2

LENGTH: 668

TYPE: PRT

ORGANISM: Homo sapiens

US-09-930-181-2

Query Match 100.0%; Score 3497; DB 4; Length 668;
Best Local Similarity 100.0%; Pred. No. 3.4e-251;
Matches 668; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy	1	MTSTGDKGGAQYVGPYRLEKTLGKGQGLVGLGVHCVTCQKVAIKIVNREKLSVVL	60
Db	1	MTSTGDKGGAQYVGPYRLEKTLGKGQGLVGLGVHCVTCQKVAIKIVNREKLSVVL	60
Qy	61	MKVEREIAILKLEHPPHVLKLDVYENKYLVLVLEHVSGLGELFDYLVKKGRLTPKEARK	120
Db	61	MKVEREIAILKLEHPPHVLKLDVYENKYLVLVLEHVSGLGELFDYLVKKGRLTPKEARK	120
Qy	121	PFROIISALDFCHSHSICHRLDKPENLLDEKKNIRADFGMASLQVDSLETSCGSPH	180
Db	121	PFROIISALDFCHSHSICHRLDKPENLLDEKKNIRADFGMASLQVDSLETSCGSPH	180
Qy	181	YACFEVIRGEKYGDKADVNSCGVILFALLVGLPFPDDNLRQLLEKVKRGVFMHPHIP	240
Db	181	YACFEVIRGEKYGDKADVNSCGVILFALLVGLPFPDDNLRQLLEKVKRGVFMHPHIP	240
Qy	241	PDCQSLLRGMIEVDAAARLTLKHIQKHIWYIGGKNEPEPEQPIPRKQVIRSLPSLEDIDP	300
Db	241	PDCQSLLRGMIEVDAAARLTLKHIQKHIWYIGGKNEPEPEQPIPRKQVIRSLPSLEDIDP	300
Qy	301	DVLDSMSLGCFRDRNKLQDLSSEENQEKMIYFLLLDRKERYPSQDEDLPPRNEIDP	360
Db	301	DVLDSMSLGCFRDRNKLQDLSSEENQEKMIYFLLLDRKERYPSQDEDLPPRNEIDP	360
Qy	361	PRKVDSPMLNRHCKRRPERKSMVLSVTDGGSVPARRAITEMAQHGQSRSSISGASSGL	420
Db	361	PRKVDSPMLNRHCKRRPERKSMVLSVTDGGSVPARRAITEMAQHGQSRSSISGASSGL	420

5962312

Qy 421 STPLSGSPRVTPHSPGRSGPLTPKGTGPHVHTPKESPAGTNPPTPPSPSPVGGVPMARLN 480
Db 421 STPLSGSPRVTPHSPGRSGPLTPKGTGPHVHTPKESPAGTNPPTPPSPSPVGGVPMARLN 480
Qy 481 SIKNSFLGSPRFRHRKLVQVTPPEMSNLTPESSPELAKKSWFGNFISLEKEQIFVVIKD 540
Db 481 SIKNSFLGSPRFRHRKLVQVTPPEMSNLTPESSPELAKKSWFGNFISLEKEQIFVVIKD 540
Qy 541 KPLSSIKADIVHAFSLTPSLSHSVISOTSPRAEYKATGGPAVQKPKVQFVDITYTEGGE 600
Db 541 KPLSSIKADIVHAFSLTPSLSHSVISOTSPRAEYKATGGPAVQKPKVQFVDITYTEGGE 600
Qy 601 AQKENGYSYVFTLLSGSPRRFRKRVETIOAQLSTHDDPPAAQHLSDTTNCMEMMTGRLS 660
Db 601 AQKENGYSYVFTLLSGSPRRFRKRVETIOAQLSTHDDPPAAQHLSDTTNCMEMMTGRLS 660
Qy 661 KCGIIPKS 668
Db 661 KCGIIPKS 668

RESULT 2
US-09-930-181-4
; Sequence 4, Application US/09930181
; Patent No. 6455292
; GENERAL INFORMATION:
; APPLICANT: Origene Technologies
; TITLE OF INVENTION: Full-Length Serine Protein Kinase in Brain and Pancreas
; FILE REFERENCE: 160 101 VI
; CURRENT APPLICATION NUMBER: US/09/930,181
; CURRENT FILING DATE: 2001-08-16
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 4
; LENGTH: 585
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-930-181-4

Query Match 86.6%; Score 3028; DB 4; Length 585;
Best Local Similarity 100.0%; Pred. No. 1.7e-216;
Matches 577; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 92 YLVLEHVS GSGELFDYL VKKGR LTPKEARK FFRQII SALDFCHSHS ICHRD LKPN LLLDE 151
Db 9 YLVLEHVS GSGELFDYL VKKGR LTPKEARK FFRQII SALDFCHSHS ICHRD LKPN LLLDE 68
Qy 152 KNRIADFGMASLQVGDLSLETSCGSPHYACPEVIRGEKYDGRKADVWSCGVILFALLV 211
Db 69 KNRIADFGMASLQVGDLSLETSCGSPHYACPEVIRGEKYDGRKADVWSCGVILFALLV 128
Qy 212 GALLPDDNNLQLLEKVKRGVFMHPHFTPPDCQSLLRGMIEVDAARLTLLEHIQKHIWYI 271
Db 129 GALLPDDNNLQLLEKVKRGVFMHPHFTPPDCQSLLRGMIEVDAARLTLLEHIQKHIWYI 188
Qy 272 GKQNEPEQPIPRKVQIRSLPSLEIDDPDVLDSMHSIGCFRDRNKLQDLLSEENQEK 331
Db 189 GKQNEPEQPIPRKVQIRSLPSLEIDDPDVLDSMHSIGCFRDRNKLQDLLSEENQEK 248
Qy 332 MIYFLLDRKERYPSQEDLPPRNEIDPPRKYDSDPMLNHRGKRPERKSMVLSVTDG 391
Db 249 MIYFLLDRKERYPSQEDLPPRNEIDPPRKYDSDPMLNHRGKRPERKSMVLSVTDG 308
Qy 392 GSPVARRAIEAQAQGRSISGASSGLSTSPSSPRVTPHPSPRGSPLTPKGTGPHVT 451
Db 309 GSPVARRAIEAQAQGRSISGASSGLSTSPSSPRVTPHPSPRGSPLTPKGTGPHVT 368
Qy 452 PKESPAGTNPPTPPSPSPVGGVPMARLNSIKNSFLGSPRFRHRKLVQVTPPEMSNLTP 511
Db 369 PKESPAGTNPPTPPSPSPVGGVPMARLNSIKNSFLGSPRFRHRKLVQVTPPEMSNLTP 428
Qy 512 SSPELAKKSWFGNFISLEKEQIFVVIKDPLSSIKADIVHAFSLTPSLSHSVISOTSFR 571

Db 429 SSPELAKKSWFGNFISLEKEQIFVVIKDPLSSIKADIVHAFSLTPSLSHSVISOTSFR 488
Qy 572 AYKATGGA VQKPKVQFVDITYTEGGEAQKENGYSYVFTLLSGSPRRFRKRVETIOA 631
Db 489 AYKATGGA VQKPKVQFVDITYTEGGEAQKENGYSYVFTLLSGSPRRFRKRVETIOA 548
Qy 632 QLLSTHDDPPAAQHLSDTTNCMEMMTGRLSKCGIIPKS 668
Db 549 QLLSTHDDPPAAQHLSDTTNCMEMMTGRLSKCGIIPKS 585

RESULT 3
US-09-930-181-17
; Sequence 17, Application US/09930181
; Patent No. 6455292
; GENERAL INFORMATION:
; APPLICANT: Origene Technologies
; TITLE OF INVENTION: Full-length Serine Protein Kinase in Brain and Pancreas
; FILE REFERENCE: 160 101 VI
; CURRENT APPLICATION NUMBER: US/09/930,181
; CURRENT FILING DATE: 2001-08-16
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 17
; LENGTH: 603
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-930-181-17

Query Match 86.2%; Score 3015; DB 4; Length 603;
Best Local Similarity 99.7%; Pred. No. 1.6e-215;
Matches 574; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 72 LIEHPVLKLDVYENKYLVLVLEHVS GSGELFDYL VKKGR LTPKEARK FFRQII SALDF 131
Db 1 LIEHPVLKLDVYENKYLVLVLEHVS GSGELFDYL VKKGR LTPKEARK FFRQII SALDF 60
Qy 132 CHSHSICHRLDKPENLLDEKKNIRIADFGMASLQVGDLSLETSCGSPHYACPEVIRGEK 191
Db 61 CHSHSICHRLDKPENLLDEKKNIRIADFGMASLQVGDLSLETSCGSPHYACPEVIRGEK 120
Qy 192 YDGRKADVWSCGVILFALLVGLPDDNNLRLLEKVKRGVFMHPHFTPPDCQSLLRGMI 251
Db 121 YDGRKADVWSCGVILFALLVGLPDDNNLRLLEKVKRGVFMHPHFTPPDCQSLLRGMS 180
Qy 252 EVDAARLTLLEHIQKHIWYIGGKNEPEQPIPRKVQIRSLPSLEIDDPDVLDSMHSIGC 311
Db 181 EVDAARLTLLEHIQKHIWYIGGKNEPEQPIPRKVQIRSLPSLEIDDPDVLDSMHSIGC 240
Qy 312 FRDRNKLQDLLSEENQEKMIYFLLDRKERYPSQEDLPPRNEIDPPRKYDSDPMLN 371
Db 241 FRDRNKLQDLLSEENQEKMIYFLLDRKERYPSQEDLPPRNEIDPPRKYDSDPMLN 300
Qy 372 RHGKRPERKSMVLSVTDGSPVPARRAIEAQAQGRSISGASSGLSTSPSSPRVT 431
Db 301 RHGKRPERKSMVLSVTDGSPVPARRAIEAQAQGRSISGASSGLSTSPSSPRVT 360
Qy 432 PHPSRGSPLTPKGTGPHVHTPKESPAGTNPPTPPSPSPVGGVPMARLNSIKNSFLGSPR 491
Db 361 PHPSRGSPLTPKGTGPHVHTPKESPAGTNPPTPPSPSPVGGVPMARLNSIKNSFLGSPR 420
Qy 492 FHRKLVQVTPPEMSNLTPESSPELAKKSWFGNFISLEKEQIFVVIKQPLSSIKADIV 551
Db 421 FHRKLVQVTPPEMSNLTPESSPELAKKSWFGNFISLEKEQIFVVIKQPLSSIKADIV 480
Qy 552 HAFSLTPSLSHSVISOTSPRAEYKATGGPAVQKPKVQFVDITYTEGGEAQKENGYSVT 611
Db 481 HAFSLTPSLSHSVISOTSPRAEYKATGGPAVQKPKVQFVDITYTEGGEAQKENGYSVT 540
Qy 612 FTLLSGSPRRFRKRVETIOAQLSTHDDPPAAQHLSD 647
Db 541 FTLLSGSPRRFRKRVETIOAQLSTHDDPPAAQHLSE 576

Db 565 PSCHSQRRGASGSIKSF--TSKVRNRLNPEESKDRVETLRRPHVVGGNDKEKEPR 622
QY 511 ESSPELAKSW-FGNFISLEKEQIFVWIKDKPLSSIKADIVHAFISLSLSHSVTSQTS 569
Db 623 EAKPRSLRTWSMKTTSSMEPNEMREIRKVLNDANSQSELHEKYML--CMHGTPGHED 680
QY 570 FRAEYATGPAVFQKPVKQFVDITVTEGGEAQKENGIVSVTFTLLSGPSRRFRKRVETI 629
Db 681 F-----VQWEMEVC-----KLPRLSLNGVRFRKISGTSMAFNKIASKI 718
QY 630 QAOL 633
Db 719 ANEL 722

RESULT 6
US-09-984-890-4
; Sequence 4, Application US/09984890
; Patent No. 6492156
; GENERAL INFORMATION:
; APPLICANT: YAN, Chunhua et al.
; TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC
; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES
; FILE REFERENCE: CLO01306
; CURRENT APPLICATION NUMBER: US/09/984,890
; CURRENT FILING DATE: 2001-10-31
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 722
; TYPE: PRT
; ORGANISM: Rattus norvegicus
US-09-984-890-4

Query Match 22.7%; Score 792.5; DB 4; Length 722;
Best Local Similarity 30.2%; Pred. No. 1.2e-50;
Matches 217; Conservative 115; Mismatches 251; Indels 135; Gaps 18;
QY 10 AQHAQVGPVRLKTKGQGTGLVKGVCVTCQKVAIKVNRKLSVLMKVEREIAI 69
Db 44 ADEQPHIGNYRLTKTGKGNFAKVKLARHILTGKEVAVKIIDKTLNSSLQKLFREVI 103
QY 70 LKLTIEHPVLKLDHVVYENKYLVLVLEHVGSGGELFDYLVKKGRLTPKEARKFRQIISAL 129
Db 104 MKVLNHNIVKLFVETETKTLVLMVYASGGEVFDYLVHGRMKEKARAKFRQIVSAV 163
QY 130 DFCHSHSICHRDLKPENLLDEKNRIADFGMASLQVGSLLTSCGSPHYACPEVIRG 189
Db 164 QYCHQKFIVHRDLKAENLLDADNMNIADFGFSNEFTFGNKLDTCGSPPYAAPLFOG 223
QY 190 EKYDGRKADVWSCGVILFALLVGLPDDNLRQLLEKVKRGVFMHPHFIPDQCSSLRG 249
Db 224 KKYDGPVDVWNLGVILTYLTVSGSLPDDGNLRELRLRVLRGKRIYFPFYSTDCENLLKK 283
QY 250 MIEYDAARRLTLEHIQHIWYIGKNEPEPEQPIPRKVQIRSLPSLEIDIPDVLDSMHSI 309
Db 284 FLILNPSKRGTELEQIMKDRWNVGHEDE-----LKPVEPLPDYK--DPRTELNVSM 335
QY 310 GCFDRNKLQDLSEENKMTYFILLDRKRYPSQDEDELPNREID----- 359
Db 336 GYTREE---IQDSLVGQRYNEVATYLLGYKSSLEGGDTITLTPRPSADLTNSSAPSPS 392
QY 360 -----PPEKVD-----PMLNRHCK-----RPE-----RKSMEVLSVTD 390
Db 393 HKVQSVSAMPKQRRSDQAVPAIPTNSVSKTQSNNAENKRPETEETGRKASTAKVP- 451
QY 391 GGSVPARRAETEMAHQGRSRSISGASGLSTPSL-----SSPRVTPH 433
Db 452 -ASPLGLDRKKTFTTSTNSVLSSTNSRNSPFLDRASLQASIQNGKSTAPQRPV 510
QY 434 PSRGSPLPTPKGTPVHTPKESPAGT-----PNTPPSSPSVGGV 473

Db 511 ASPSAHNSSSSGADRT--NFRGVSSRSTFHAGQLQVRDQQLPRGVTTPASPSGHSQ 568
QY 474 PWRARLNIKSNFLSGPRHRRKLVQV-TPREMSNLTP-----ESSPEL 516
Db 569 GRRGASGSIKSF--TSKVRNRLNPEESKDRVETLRRPHVVGGGTDKEKEFRKAPRS 626
QY 517 AKSW-FGNFISLEKEQIFVWIKDKPLSSIKADIVHAFISLSLSHSVTSQTSFRAEYK 575
Db 627 LRFTWSMKTTSMENEMREIRKVLNDANSQSELHEKYML--CVHGTGPHENP----- 679
QY 576 ATGGAQVQKPVKQFVDITVTEGGEAQKENGIVSVTFTLLSGPSRRFRKRVETIQAOL 633
Db 680 -----VQWEMEVC-----KLPRLSLNGVRFRKISGTSMAFNKIASKIANEL 720

RESULT 7
US-08-817-832B-32
; Sequence 32, Application US/08817832B
; Patent No. 6579691
; GENERAL INFORMATION:
; APPLICANT: MANDELKOW, Eckhard, et al.
; TITLE OF INVENTION: NO. 6579691el Protein Kinase (NPK-110)
; NUMBER OF SEQUENCES: 32
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
; STREET: 233 S. Wacker Drive, 6300 Sears Tower
; CITY: Chicago
; STATE: Illinois
; COUNTRY: US
; ZIP: 60606-6402
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/817,832B
; FILING DATE: 28-APR-1997
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: NO PCT/EP95/04258
; FILING DATE: 30-OCT-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: EP 94 11 7122.5
; FILING DATE: 28-OCT-1994
; INFORMATION FOR SEQ ID NO: 32:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 722 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-817-832B-32

Query Match 22.6%; Score 790.5; DB 4; Length 722;
Best Local Similarity 30.2%; Pred. No. 1.7e-50;
Matches 217; Conservative 115; Mismatches 251; Indels 135; Gaps 18;
QY 10 AQHAQVGPVRLKTKGQGTGLVKGVCVTCQKVAIKVNRKLSVLMKVEREIAI 69
Db 44 ADEQPHIGNYRLTKTGKGNFAKVKLARHILTGKEVAVKIIDKTLNSSLQKLFREVI 103
QY 70 LKLTIEHPVLKLDHVVYENKYLVLVLEHVGSGGELFDYLVKKGRLTPKEARKFRQIISAL 129
Db 104 MKVLNHNIVKLFVETETKTLVLMVYASGGEVFDYLVHGRMKEKARAKFRQIVSAV 163
QY 130 DFCHSHSICHRDLKPENLLDEKNRIADFGMASLQVGSLLTSCGSPHYACPEVIRG 189
Db 164 QYCHHKKFIVHRDLKAENLLDADNMNIADFGFSNEFTFGNKLDTCGSPPYAAPLFOG 223
QY 190 EKYDGRKADVWSCGVILFALLVGLPDDNLRQLLEKVKRGVFMHPHFIPDQCSSLRG 249
Db 224 KKYDGPVDVWNLGVILTYLTVSGSLPDDGNLRELRLRVLRGKRIYFPFYSTDCENLLKK 283

APPLICANT: Bandman, Olga
APPLICANT: Molteni, Angela
APPLICANT: Magnaghi, Paola
APPLICANT: Bosotti, Roberta
APPLICANT: Scacheri, Emanuela
APPLICANT: Isacchi, Antonella
APPLICANT: Hodgson, Dave
TITLE OF INVENTION: HUMAN NIM1 KINASE
FILE REFERENCE: PC-0009 US
CURRENT APPLICATION NUMBER: US/09/523,849
CURRENT FILING DATE: 2000-03-13
NUMBER OF SEQ ID NOS: 39
SOFTWARE: PERL Program
SEQ ID NO 34
LENGTH: 776
TYPE: PRT
ORGANISM: Rattus norvegicus
FEATURE:
NAME/KEY: misc.feature
OTHER INFORMATION: GenBank Accession No. 6458561 g5672676
US-09-523-849-34

Query Match 21.4%; Score 749; DB 4; Length 776;
Best Local Similarity 31.5%; Pred. No. 2.2e-47;
Matches 214; Conservative 87; Mismatches 223; Indels 156; Gaps 20;
4 TGKGGGAHQAGYGVPRLEKTLGKGTGLKGLVHCVTCKVAIKIVNRKLSVLMKV 63
12 TGTGGQKPLRVGYDVERTLGKGNFVAVKLAHRVTKTQVAIKIDKTLDSNNLEKI 71
64 ERETAIKLIEHPVHLKLDVYENKYLVLVLEHVSQGLFDYLVKKGRLTPKARKFR 123
72 YREVLAKLNLHPNIIKLYQVWETKMLYIVTEPAKNGEMFDYLTNGHLSSENEARKFW 131
124 QIISALDFCHSHSICHRLDKENLLLEKKNIRIADFGMASLQVDSLLTSCSPHYAC 183
132 QILSAVEYCHNHIIVHRDLKTNELLDGNMDIKLADFGNFYKPGFPLTWCSPYAA 191
184 PEVIRGEKYDGRKADVMSGCVILFALLVAGLPFDNDNLRLLEKVKRGVFMHPHFPDPC 243
192 PEVEGKEYEGPQLDIWSGLVVLVCGSLPFDGPNLPTLRQVLRGRFRIFPFMSQDC 251
244 QSLRGMIEVDAAARLTLEHKIHWYIGGKNEPEP---EPIPRKVOIRSLPSLEIDP 300
252 ETLLRRLVMDPAKRITIAQIRHRW-----QADPTLLQDDPAFSWQGYTSLG DYNE 306
301 DVLSMHSGLCFRDRNKLQDLSSEENQERKMYFLILDR-----KERYPSQDEDLPPR- 355
307 QVLGIMQALGI--DRQRTVBSLQSSYNHFAAIYLLLERLRHRSTQPSRATPAPARQ 364
356 -----NEI---DPPRKRVDSPP-----LNRHGKRRP----- 378
365 POLRNSDLSLSEVQEILPCDPFRPSLLCPOQALQSVLQAEICDLHSLQPLFPFLD 424
379 -----ERKMEVLVTDGSPVPAARRAIMAGHQRSSISGASSGLST-----S 423
425 TNCSGVFRHRSISPSLLDTAISEARQGSLEBEEQVEPLPGSTCRRHTLAESVTHFS 484
424 PLSSPRVTPHSPRGSPLTPKGT-----VHTKESPA-----TPNTPPSPSVGG 472
485 PLNPPCIIVSSSAVSP---SEGSSSDCLFSASEGPAIGGLGATPGLLGTSPP--- 537
473 VPMRLNLSIKNSFLGSPFRHRRKLQVTPREMSN-----LTPESPE----- 515
538 -----VRLAS-----PFLGS-----QSATPVLSQAGLGATVLPVFOEGRRASDTSLT 582
516 -----LAKSWFGNFIISLEKEQIFVVIKOKPLSSIKADIVHAPLSIPSLSHVIS 566
583 QGLKAFRQOLKRNARTKGLGLNK-----IKGLARQV-C 615
567 QTSFRAEYKATGGPAVFOKP 586
616 QSSIRG-----SRGGMSTFHTP 632

RESULT 12

US-08-817-832B-31
Sequence 31, Application US/08817832B
Patent No. 6579691
GENERAL INFORMATION:
APPLICANT: MANDELKOW, Eckhard, et al.
TITLE OF INVENTION: No. 6579691el Protein Kinase (NPK-110)
NUMBER OF SEQUENCES: 32
CORRESPONDENCE ADDRESS:
ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
STREET: 233 S. Wacker Drive, 6300 Sears Tower
CITY: Chicago
STATE: Illinois
COUNTRY: US
ZIP: 60606-6402
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30 (EPO)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/817,832B
FILING DATE: 28-APR-1997
PRIOR APPLICATION DATA:
APPLICATION NUMBER: WO PCT/EP95/04258
FILING DATE: 30-OCT-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: EP 94 11 7122.5
FILING DATE: 28-OCT-1994
INFORMATION FOR SEQ ID NO: 31:
SEQUENCE CHARACTERISTICS:
LENGTH: 779 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-817-832B-31

Query Match 21.4%; Score 749; DB 4; Length 779;
Best Local Similarity 32.4%; Pred. No. 2.2e-47;
Matches 192; Conservative 94; Mismatches 186; Indels 120; Gaps 15;
15 YGPPYLEKTLGKGTGLKGLVHCVTCKVAIKIVNRKLSVLMKVRETAILKLE 74
42 HIGNYRLQTKIGKGNFVAVKLAHRVTGREVAVKIIDKTLNFTSLQKLFREVRIMKILN 101
75 HPHVLKLDVYENKYLVLVLEHVSQGLFDYLVKKGRLTPKARKFRQIISALDFCHS 134
102 HPNIVKLFVEIETKTLVMEYASGGEVFDYLVAHGRMKERAKFRQIVSAVQYCHQ 161
135 HSICHRLDKENLLLEKKNIRIADFGMASLQVDSLLTSCSPHYACPEVIRGEKYD 194
162 KCIVHRDLKAENLLLDADNMNIKIADFGPSNEFTVGNKLDITFCGSPPYAAPELPQGRKYD 221
195 RKADVMSGCVILFALLVAGLPFDNDNLRLLEK-VKRGVFMHPHFPDCCQSLRGMIEV 253
222 PEVDVMSGLVILTVSGSLPFDGQNLKELRERSCLRGKRVFPYVWSTDCENLLKLLVL 281
254 DAARRLTLEHKIHWYIGGKNEPEPEQPIPRKVOIRSLPSLEIDPDLVDSMHSGLCFR 313
282 NPIKRGSLQIKMDRWNVNGHEBEE-----LKPSEPELDLNDAKRIDIMVTMGFAR 333
314 DRNKLQDLSSEENQERKMYFLILDRK----- 341
334 DE---INDALVSKQYDEVMTATYILGRKPPFEGGESLSGSLCQSRPSSDLNNSTLQS 390
342 -----ERYPSQDEDLPP-----RNEIDPPRKRV-DSPMLNRHGK 375
391 PAHLKVQRTISANQKORRRFSDHAGSIPPAVSVYTKRQANSVESEKQEWKDKTARLGS 450
376 RR-----PERKSMEV---LSVTGGSPVPARRAI-----EMAQHGQRS 410

```
Db 451 TTGSKSEVTASPLVGDRKSSAGSNVYSGSMTRRTNYTCERSTDRYAALQNGRDS 510
Qy 411 RSTSGASSGLSTS-----PLSSPRVTPHSPRGSP-----LPTPK-GTPVHTPKESPA 457
Db 511 SLTMSASSMSSTGSTVASAGSPARPRHQKMSSTSGHPKIVTLPTTKDGSEAYRP----- 565
Qy 458 GTENPTPPSPSPVGGV----PWEARL---NSIKNSFLGSPRFHRRKLQVPTP 502
Db 566 GTAQRPAAAPSASHSISASTPDTRTPRGSSRSSTHGEQLRERRSAAVSGP 617

RESULT 13
US-08-557-006C-40
; Sequence 40, Application US/08557006C
; Patent No. 6258547
; GENERAL INFORMATION:
; APPLICANT: Beri, Rajindar K.
; APPLICANT: Carling, David
; APPLICANT: Forster, Robert A.
; TITLE OF INVENTION: NUCLEIC ACID ENCODING AMP-ACTIVATED PROTEIN KINASE
; FILE REFERENCE: NGAP/PHM37588/UST
; CURRENT APPLICATION NUMBER: US/08/557,006C
; CURRENT FILING DATE: 1996-03-06
; PRIOR APPLICATION NUMBER: PCT/GB94/01093
; PRIOR FILING DATE: 1994-05-20
; PRIOR APPLICATION NUMBER: GB 9310489.1
; PRIOR FILING DATE: 1993-05-21
; PRIOR APPLICATION NUMBER: GB 9318010.7
; PRIOR FILING DATE: 1993-08-31
; NUMBER OF SEQ ID NOS: 44
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 40
; LENGTH: 552
; TYPE: PRT
; ORGANISM: Rat
; FEATURE:
; NAME/KEY: gene
; LOCATION: (1)..(1747)
; OTHER INFORMATION: Full length cDNA sequence fragment of Human AMPK -
; OTHER INFORMATION: fragment begins at nucleotide 24 and ends with
; OTHER INFORMATION: nucleotide 1765
US-08-557-006C-40

Query Match 20.9%; Score 729.5; DB 3; Length 552;
Best Local Similarity 29.8%; Pred. No. 4e-46;
Matches 197; Conservative 103; Mismatches 217; Indels 143; Gaps 22;

Qy 16 VGPYRLKTLGKGTGLVKLVGHVCTCKVAIKIVNREKL-SESVLMMKVEREAILKLE 74
Db 13 IGHVYLGDTLGGTGGKVGKIGEHQLTGHKAVKILNRQKIRSLDVGVGKIKREIQNLKLF 72
Qy 75 HPHVLKLDHVENKKYLVLVLEHVSGLDFYLVKKGRLTPKEARKFPRQIISALDFCHS 134
Db 73 HPHIILKQVISTPTDFFMVVEYSGSELDYICKHGRVEVEARRLFFQQLSLSAVDYCHR 132
Qy 135 HSIHRDLKPENLLDEKNRIADFGMASLQVGSLSLETSCGSPHYACPEVIRGEKYDG 194
Db 133 HMVVRDLKPENLLDAQMNKADFGLSNMMSDGEFLRTSCGSPNYAAPEVIRGLYAG 192
Qy 195 KADVWVSCGVILFALLVGLPDDDNRLQLLEKVRGVFHPHPIPPDCQSLLRGMIEVD 254
Db 193 PEVDIWSGVILYALLCGTLFPDDHVPITLTKIRGGVFYIPEVLNRSIATLLMHMLQVD 252
Qy 255 AARRLTLEHIOKHIVIGGKNEPEQPIPRKVQIRSLPSLEDDIPDVLDS-----MHS 308
Db 253 PLKRATIKDIRHEWF-----KQDLPVSL-----FPEDPSYDANVIDEAVEKCEK 299
Qy 309 LGCFDRNKLQLDLS--EENQKRMFYILLDRKERYPSQDE----DLPPRNEIDPPRK 363
Db 300 FEC--TESEVMNSLSGDPQDLAVAYHLIDNR-RINNQASEFYLAASSPTG----- 349
Qy 364 RVDSPMLNRHGKRRPERKSMVELSVTDGGSFVPARRAIEAQAHQGRSRSISGASSGLSTS 423
```

```
Db 350 -----SPMD-----DMAMH----- 358
Qy 424 PLSSPRVTPHSPRGSP--L--PTPKGTPVHTPKESPACTPNPTPPSSPSVGGVPMPEARLNS 481
Db 359 --IPGLKPHPE-RMPPLIADSPKA-----RCPDLALNTTKPKSLAVKAKWHLGIRS 408
Qy 482 IKNSFLGSPRFHRRKLQVPTPEMSNL-----TPESSPELAKKSWFCNFIISLEKEEQIFV 537
Db 409 QSKPYDIAEVYRAMKQLDFEKVKNVAYHLRVRKNPVT-----GNVYKMSL--QLYLV 460
Qy 538 IKDKPLSLSTKADIVHAFISIPSLSHSVISQTSFRAYKATGGPAVFPQKPVQVDITYTE 597
Db 461 -----DNRSYLLDFKSIDDEWVEQRSGSSTPQSCSAAGLHRP-RSSVDSSTAE 508
Qy 598 GGEAQKENGIIYSVTFTLLSGPSRRFKRVVETIIQAQLLSTHDPAAQHLSDDTN-CWEMMT 656
Db 509 NHS-----LSG-----SLTGSUTGSTLSASPRLGSHMTDFFEMCASLIT 548

RESULT 14
US-08-557-006C-43
; Sequence 43, Application US/08557006C
; Patent No. 6258547
; GENERAL INFORMATION:
; APPLICANT: Beri, Rajindar K.
; APPLICANT: Carling, David
; APPLICANT: Forster, Robert A.
; TITLE OF INVENTION: NUCLEIC ACID ENCODING AMP-ACTIVATED PROTEIN KINASE
; FILE REFERENCE: NGAP/PHM37588/UST
; CURRENT APPLICATION NUMBER: US/08/557,006C
; CURRENT FILING DATE: 1996-03-06
; PRIOR APPLICATION NUMBER: PCT/GB94/01093
; PRIOR FILING DATE: 1994-05-20
; PRIOR APPLICATION NUMBER: GB 9310489.1
; PRIOR FILING DATE: 1993-05-21
; PRIOR APPLICATION NUMBER: GB 9318010.7
; PRIOR FILING DATE: 1993-08-31
; NUMBER OF SEQ ID NOS: 44
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 43
; LENGTH: 633
; TYPE: PRT
; ORGANISM: Yeast
; FEATURE:
; NAME/KEY: PEPTIDE
; LOCATION: (1)..(633)
; OTHER INFORMATION: Yeast SNF1 polypeptide
US-08-557-006C-43

Query Match 20.6%; Score 720.5; DB 3; Length 633;
Best Local Similarity 33.6%; Pred. No. 2.2e-45;
Matches 158; Conservative 109; Mismatches 152; Indels 51; Gaps 11;

Qy 15 YGYPYRLKTLGKGTGLVKLVGHVCTCKVAIKIVNREKLSVLM-KVEREIALKLI 73
Db 51 HIGNYQIVTKLGEISGFGKVLAYHTTTGQKVALKIINKKVLAKSDMQGRIEISYRL 110
Qy 74 EHPVLKLDHVENKKYLVLVLEHVSGLDFYLVKKGRLTPKEARKFPRQIISALDFCH 133
Db 111 RHPHILKLDHVENKKYLVLVLEHVSGLDFYLVKKGRLTPKEARKFPRQIISALDFCH 169
Qy 134 HSIHRDLKPENLLDEKNRIADFGMASLQVGSLSLETSCGSPHYACPEVIRGEKYD 193
Db 170 RHKIVHRDLKPENLLDEHNLNVIADFGLSNIMTDGNFLKTCGSPNYAAPEVIRGLYA 229
Qy 194 GRADVWVSCGVILFALLVGLPDDDNRLQLLEKVRGVFHPHPIPPDCQSLLRGMIEV 253
Db 230 GPEVDIWSGVILYALLCGTLFPDDHVPITLTKIRSGVYVTLTKPLSPGAGLIRKMLIV 289
Qy 254 DAARRLTLEHIOKHIVIGG-----KNEPEPEQPIPRKVQIRSLPSLEDDIPDVL 303
Db 290 NPLNRISIHIMQDDWFKVDLPYLLPDLKPHPEENENNDSKCGSSPDNDIEDNLV 349
```

Qy 304 DSMHSLGCFDRNRKLLQDLSEEE-----NOKMIYFLLDRKERYPSQE-----DEDLPP 354
Db 350 NILSSTWGY-EKDEIYESLESSEDTPAFNEIRDAYMLIKENKSLIKOMKANKSVSDLDLT 408
Qy 355 RNEIDPP-----RKRVDSFMLNRHGKRRERKSRMEVLSVTDGSGSPVAPARRAEMA 404
Db 409 FLSOSPPTFOQSKSHQSKQSDVHTAKQHARM-----ASAITQORTYHQS 454
Qy 405 QHGORSRSISGASSGLSTPLSSPRV--TPHSPRGSPLPPTKGPVHTPK 453
Db 455 PFMDQYKE---EDSTVILPILSLPQIHANMLAQGSPAAS-KISPLVTKK 500

RESULT 15

US-09-101-146-1
; Sequence 1, Application US/09101146
; Patent No. 6124125
; GENERAL INFORMATION:
; APPLICANT: Dartmouth College, St. Vincents Institute of
; APPLICANT: Medical Research, Kemp et al.
; TITLE OF INVENTION: No. 6124125el AMP Activated Protein Kinase
; NUMBER OF SEQUENCES: 64
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Jane Massey Licata, Esq.
; STREET: 66 E. Main Street
; CITY: Marlton
; STATE: NJ
; COUNTRY: USA
; ZIP: 08053

COMPUTER READABLE FORM:
MEDIUM TYPE: DISKETTE, 3.5 INCH, 1.44 Mb STORAGE
COMPUTER: IBM PC

OPERATING SYSTEM: WINDOWS 95
SOFTWARE: WORDPERFECT 6.0 FOR WINDOWS
CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/101.146
FILING DATE: October 7, 1998
CLASSIFICATION: 435

PRIOR APPLICATION DATA:
APPLICATION NUMBER: PN7450

FILING DATE: 8 JAN 1996
ATTORNEY/AGENT INFORMATION:

NAME: Jane Massey Licata
REGISTRATION NUMBER: 32,257
REFERENCE/DOCKET NUMBER: DC-0050

TELECOMMUNICATION INFORMATION:
TELEPHONE: (856) 810-1515

TELEFAX: (856) 810-1454
INFORMATION FOR SEQ ID NO: 1:

SEQUENCE CHARACTERISTICS:
LENGTH: 345

TYPE: Amino acid
TOPOLOGY: Linear

US-09-101-146-1

Query Match 20.5%; Score 718.5; DB 3; Length 345;
Best Local Similarity 42.5%; Pred. No. 1.4e-45;
Matches 150; Conservative 63; Mismatches 107; Indels 33; Gaps 7;

Qy 16 VGPYRLKTLGKQGTGLVGLVCHVTCQKVAIKIVNREKLSVLMKVEREIALKLIH 75
Db 12 IGHVILGDTLGVGTGKVGKVGKHELTGHKAVKILNRQKIRLDVVVGKIRREIQNLKLF 71
Qy 76 PHVLKLDHVVYENKYLVLVLESHVSGGELFDYLVKKGLTTPKEARKFROIISALDFCHSH 135
Db 72 PHILKLVQVISTPDSIDFMVMEYVSGGELFDYICKNGRLDEKESRRLFOQLSGVDYCHRH 131
Qy 136 SIHRDLKPNLLDEKKNIRIADFGMASLQVGSLSLETSCSPHYACPEVIRGEKYDGR 195
Db 132 MVVHRDLKPNVLLDAHNNAKIAFDGLSNMWSGDEFLTSCSPNYAAPEVISGRLYAGP 191
Qy 196 KADVWSCGVILFALLVGLALPFDDNNLRQLLEKVKRGVFMHPHFIPTPCQSLLRGMIEYDA 255

Db 192 EVDIWSGVLIALCGTLFPDDHVPITLTKKICDGFYTPQYLNPSVISLLKXHLQVDP 251
Qy 256 ARELTLEHIOKHIWYIGGKNEPEPEQPIPRKVQIRSLPSLED--IDPDVLSM----- 306
Db 252 MKGATIKDIREHEWF-----KQDLP-KYLFPEDPSPYSSTMIIDDEALKEVCEKEPCS 301
Qy 307 --HSLGCFDRNRKLLQDLSEEEENQEKMIYFLLDRKERYPSQEDDL---PP 354
Db 302 EESVLSCLYNRNH--QDPLA-----VAYHLIIDNRIRIMNEAKDFYLATSPP 345

Search completed: November 26, 2003, 12:03:21
Job time : 18.7053 secs